



Universidad
Zaragoza

Trabajo Fin de Grado

Estudio de técnicas de Machine Learning en el proceso de fabricación de rodamientos

Study of Machine Learning techniques in the manufacturing
process of bearings

Autor

Diego Bazo Ezquerro

Grado en Ingeniería de Tecnologías Industriales

Director

Juan José Aguilar Martín

Dpto. de Ingeniería de Diseño y Fabricación



**Escuela de
Ingeniería y Arquitectura**
Universidad Zaragoza

2020

Tabla de contenido

Tabla de contenido.....	1
1. Introducción	1
1.1. Objetivo	1
1.2. Estructura	3
1.3. Machine Learning y su aplicación	3
2. Ajuste modelo simple.....	5
2.1. Variabilidad rectangular	5
2.1.1. Aro	5
2.1.1.1. Linear regression	8
2.1.1.2. Complex tree	9
2.1.1.3. Simple y medium tree	10
2.1.1.4. Cribado de factores	10
2.1.1.5. Introducción de ruido.....	11
2.1.2. Cono A y cono B	12
2.1.2.1. Linear regression	13
2.1.3. Holgura	13
2.1.3.1. Cribado de factores	15
2.1.3.2. Introducción de ruido.....	16
2.2. Variabilidad normal.....	16
2.2.1. Aro	17
2.2.1.1. Cribado de factores	18
2.2.1.2. Introducción de ruido.....	18
2.2.2. Holgura	21
2.2.2.1. Cribado de factores	23
3. Ajuste modelo complejo	23
3.1. Variabilidad rectangular	24
3.1.1. Aro	24
3.1.1.1. Cribado de factores	26
3.1.1.2. Introducción de ruido.....	27
3.1.2. Holgura	30
3.1.2.1. Cribado de factores	31
3.1.2.2. Introducción de ruido.....	32
3.2. Variabilidad normal.....	33
3.2.1. Aro	33

3.2.1.1. Cribado de factores	34
3.2.1.2. Introducción de ruido.....	35
3.2.2. Holgura	35
3.2.2.1. Cribado de factores	37
3.2.2.2. Introducción del ruido	38
4. Coste computacional.....	38
5. Conclusiones y trabajo futuro	39
Anexo 1. Código generación de datos inicial	41
Anexo 2. Código de generación mejorado	44
Anexo 3. Análisis con Regression Learner de Matlab	46
Bibliografía	59

Índice de figuras

Figura 1. Estudio estadístico Media	1
Figura 2. Estudio estadístico Recorrido	1
Figura 3. Plano rodamiento	3
Figura 4. Diámetro frente a n	6
Figura 5. Diámetro frente a T	6
Figura 6. Diámetro frente a incertidumbre	7
Figura 7. Diámetro frente a E15D1	7
Figura 8. Diámetro frente a vibraciones	8
Figura 9. Errores de los modelos	8
Figura 10. Respuesta predicha frente a real	8
Figura 11. Ocurrencia de valores residuales	9
Figura 12. Respuesta predicha frente a real (2)	9
Figura 13. Ocurrencia de valores residuales (2)	10
Figura 14. Ocurrencia de valores residuales sin considerar la incertidumbre	11
Figura 15. Representación RMSE frente a porcentaje de ruido	12
Figura 16. Errores de los modelos (2)	12
Figura 17. Respuesta predicha frente a real (3)	13
Figura 18. Holgura frente a n	13
Figura 19. Holgura frente a T	14
Figura 20. Holgura frente a incertidumbre	14
Figura 21. Holgura frente a consigna palpador	14
Figura 22. Holgura frente a vibración	15
Figura 23. RMSE considerando todas las variables	15
Figura 24. RMSE considerando incertidumbre del palpador y los tres diámetros	15
Figura 25. RMSE considerando incertidumbre, diámetros y una temperatura de cada pieza ..	15
Figura 26. RMSE considerando incertidumbre, diámetros y dos temperaturas de dos piezas ..	15
Figura 27. RMSE considerando incertidumbre, diámetros, una temperatura de cada pieza y las tres X	16
Figura 28. Representación RMSE frente a porcentaje de ruido (2)	16
Figura 29. Diámetro frente a n (2)	17

Figura 30. Diámetro frente a T (2)	17
Figura 31. Diámetro frente a incertidumbre (2)	17
Figura 32. Diámetro frente a vibraciones (2)	18
Figura 33. Ocurrencia de valores residuales sin considerar la incertidumbre (2)	18
Figura 34. Representación RMSE frente a porcentaje de ruido (3)	19
Figura 35. RMSE para todos los modelos de Regression Learner	20
Figura 36. RMSE para todos los modelos de Regression Learner (2)	21
Figura 37. Holgura frente a n (2)	21
Figura 38. Holgura frente a T (2)	22
Figura 39. Holgura frente a incertidumbre (2)	22
Figura 40. Holgura frente a consigna del palpador (2)	22
Figura 41. Holgura frente a vibración (2)	23
Figura 42. RMSE considerando todas las variables (2)	23
Figura 43. RMSE considerando incertidumbre, diámetros, una temperatura de cada pieza y las tres X (2)	23
Figura 44. RMSE considerando incertidumbre, diámetros y dos temperaturas de dos piezas (2)	23
Figura 45. Diámetro frente a n (3)	24
Figura 46. Diámetro frente a T (3)	25
Figura 47. Diámetro frente a incertidumbre (3)	25
Figura 48. Diámetro frente a E15D1 (2)	26
Figura 49. RMSE considerando todas las variables (3)	26
Figura 50. RMSE considerando una temperatura y la incertidumbre	26
Figura 51. RMSE considerando todas las temperaturas y la incertidumbre	26
Figura 52. RMSE considerando la incertidumbre y todas las temperaturas menos Ta	27
Figura 53. RMSE considerando la incertidumbre y todas las temperaturas menos Ta y Tt	27
Figura 54. Representación RMSE frente a porcentaje de ruido (4)	27
Figura 55. RMSE para todos los modelos de Regression Learner (3)	28
Figura 56. RMSE para todos los modelos de Regression Learner (4)	29
Figura 57. Holgura frente a n (3)	30
Figura 58. Holgura frente a T (3)	30
Figura 59. Holgura frente a incertidumbre (3)	30

Figura 60. Holgura frente a consigna palpador (3)	31
Figura 61. Holgura frente a ruido	31
Figura 62. RMSE considerando todas las variables (4)	31
Figura 63. RMSE considerando incertidumbre, diámetros, una temperatura de cada pieza y las tres X (3)	31
Figura 64. RMSE considerando incertidumbre, diámetros, todas las temperaturas menos Ta y las tres X	32
Figura 65. Representación RMSE frente a porcentaje de ruido (5)	32
Figura 66. Diámetro frente a n (4)	33
Figura 67. Diámetro frente a T (4)	33
Figura 68. Diámetro frente a incertidumbre (4)	34
Figura 69. Diámetro frente a vibraciones (3)	34
Figura 70. RMSE considerando todas las variables (5)	34
Figura 71. RMSE considerando la incertidumbre y todas las temperaturas menos Ta (2)	34
Figura 72. Representación RMSE frente a porcentaje de ruido (6)	35
Figura 73. Holgura frente a n (4)	35
Figura 74. Holgura frente a T (4)	36
Figura 75. Holgura frente a incertidumbre (4)	36
Figura 76. Holgura frente a consigna palpador (4)	37
Figura 77. Holgura frente a vibración (3)	37
Figura 78. RMSE considerando todas las variables (6)	37
Figura 79. RMSE considerando incertidumbre, diámetros, todas las temperaturas menos Ta y las tres X (2)	37
Figura 80. Representación RMSE frente a porcentaje de ruido (7)	38

Índice Tablas

Tabla 1. Variables y salida del Aro	2
Tabla 2. Variables y salida del Cono A	2
Tabla 3. Variables y salida del Cono B	2
Tabla 4. RMSE frente a porcentaje de ruido	11
Tabla 5. RMSE frente a porcentaje de ruido (2)	16
Tabla 6. RMSE frente a porcentaje de ruido (3)	19
Tabla 7. RMSE frente a porcentaje de ruido (4)	27
Tabla 8. RMSE frente a porcentaje de ruido (5)	32
Tabla 9. RMSE frente a porcentaje de ruido (6)	35
Tabla 10. RMSE frente a porcentaje de ruido (7)	38
Tabla 11. Tiempos computación	39
Tabla 12. Tiempos para diferente número de variables	39

1. Introducción

1.1. Objetivo

El objetivo es el estudio de la implementación de un modelo para poder ser capaces de predecir las dimensiones finales de un rodamiento mediante la sensorización de multitud de variables de fabricación. Este método supondría una sustancial mejora con respecto al método anterior en el que cada pieza se controla midiendo solo algunas características importantes después de su rectificado. De esta forma, mientras la pieza está siendo fabricada se procesará la información disponible y heterogénea con modelos que han sido previamente entrenados para controlar el proceso de forma óptima.

En este trabajo se abordará el estudio de la influencia de las diferentes variables en el resultado final de cada componente del rodamiento así como de la holgura total, medida crítica en el proceso de fabricación. También se estudiarán las soluciones a implementar en el proceso descartando los métodos estadísticos y proponiendo métodos de Machine Learning, mediante el uso de Matlab, así como el análisis de su robustez frente a la inclusión de ruido. Dicho estudio se realizará sobre los datos sintéticos obtenidos de diferentes modelos que simulan la recogida de datos durante un año de producción.

Se va a estudiar como poder llevar un control dimensional sobre la fabricación de rodamientos. Para comenzar se ha comprobado si la evolución de las dimensiones de los rodamientos fabricados siguen una distribución que se pueda abordar desde el punto de vista estadístico. Para ello se ha realizado un estudio estadístico (SPC) a lo largo de un día.

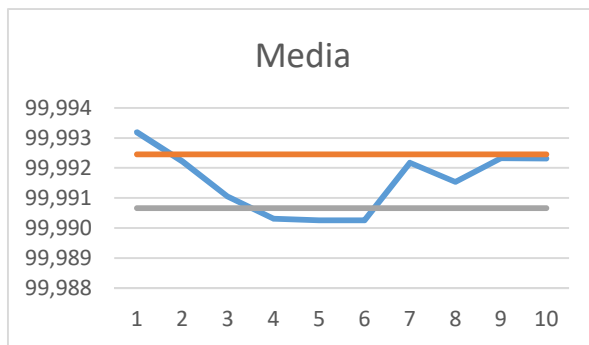


Figura 1. Estudio estadístico Media

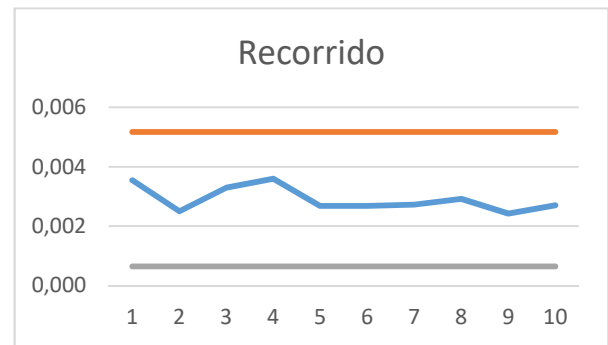


Figura 2. Estudio estadístico Recorrido

Como se puede observar con los datos de un día la evolución de las dimensiones no sigue una distribución que permita utilizar los clásicos indicadores de capacidad de proceso mediante un control estadístico. Los procesos de fabricación tienen mayor viabilidad de la que sería necesaria para ajustarse a las tolerancias exigidas. Por ello se ha incluido al final de las líneas de producción un supermercado en el que se van almacenando los componentes y se organizan por familias. La elección de los componentes de cada rodamiento se realiza de las familias compatibles para garantizar las tolerancias finales. Por todo ello se intentará explicar la variabilidad de los procesos mediante modelos de *Machine Learning*.

Para ello se dispone de las siguientes variables medibles en el proceso de fabricación:

ARO	
A1.6.Ta	T Ambiente
A1.6.V	Vibraciones
A1.6.R	Ruido
A1.6.X	Consigna Palpador
A1.6.EX	Error Consigna Palpador
A1.6.I	Incertidumbre Palpador
A1.6.Tt	Taladrina
A1.6.Tx	Eje X
A1.6.P	Consumo Potencia
A1.6.Tp	Pieza
A1.6.Tp	Pieza
A1.5.D1	
A1.6.D1	

Tabla 1. Variables y salida del Aro

Cono A	
A2.3.Ta	T Ambiente
A2.3.V	Vibraciones
A2.3.R	Ruido
A2.3.X	Consigna Palpador
A2.3.EX	Error Consigna Palpador
A2.3.D	Diamantado
A2.3.Tt	Taladrina
A2.3.Tx	Eje X
A2.3.P	Consumo Potencia
A2.3.Tp	Pieza
A2.3.Tp	Pieza
A2.2.D1	
A2.3.D1	

Tabla 2. Variables y salida del Cono A

Cono B	
A3.3.Ta	T Ambiente
A3.3.V	Vibraciones
A3.3.R	Ruido
A3.3.X	Consigna Palpador
A3.3.EX	Error Consigna Palpador
A3.3.D	Diamantado
A3.3.Tt	Taladrina
A3.3.Tx	Eje X
A3.3.P	Consumo Potencia
A3.3.Tp	Pieza
A3.3.Tp	Pieza
A3.2.D1	
A3.3.D1	

Tabla 3. Variables y salida del Cono B

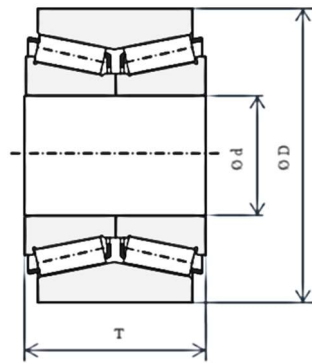


Figura 3. Plano rodamiento

El objetivo es predecir las dimensiones de los diámetros de los tres componentes, en función de las variables de fabricación, para poder predecir la cota crítica de T, de esta forma se será capaz de ordenar a un rectificador la corrección de las cotas ajustando la dimensión de T dentro de la tolerancia requerida de una forma más efectiva y previa a la fabricación.

Inicialmente no se dispone de los datos reales de las variables, por lo que se simularán los datos mediante la incursión de distintas variabilidades y se generarán los diámetros totales de la holgura mediante dos modelos, uno más simple y otro más complejo.

La generación de los datos se llevará a cabo mediante código desarrollado en Matlab. Además a través de la aplicación interna de Matlab especializada en el Machine Learning llamada 'Regression Learner' se procederá a analizar los datos obtenidos y comprobar que método matemático funcionará mejor para poder predecir los resultados.

1.2. Estructura

- Estudio de los procesos de fabricación de rodamientos y de las técnicas de Machine Learning
- Generación de datos sintéticos. Estudio de diferentes modelos
- Distribución estadística de las variables: rectangular y gaussiana
- Adición de ruido y variables múltiples
- Análisis de resultados
- Conclusiones

1.3. Machine Learning y su aplicación

Las técnicas de Machine Learning están sufriendo, como el resto de técnicas de Inteligencia Artificial, un desarrollo muy importante en los últimos años permitiendo de esta forma la posibilidad de que cada vez se les pueda sacar un mayor partido. El Machine Learning constituye una técnica por la que a través de la toma masiva de datos un sistema computacional pueda aprender y ser capaz de identificar patrones y deducir reglas autónomamente para posteriormente dar una respuesta ante problemas totalmente nuevos. Este tipo de técnicas no son nuevas y llevan décadas en la mente de las personas pero la necesidad de grandes cantidades de datos y el ser muy costosas computacionalmente han limitado su aplicación

generalizada hasta el gran desarrollo en los ordenadores y el Big Data producido a finales del siglo XX.

Dentro del Machine Learning existen dos ramas según el tipo de parámetros que se disponen inicialmente. El Aprendizaje Supervisado es aquel en el que se dispone de una serie de soluciones conocidas y de sus variables o parámetros que forjaran el aprendizaje del sistema computacional para posteriormente ser capaz de deducir y pronosticar la solución a través de solo los parámetros medidos. El Aprendizaje No Supervisado por el contrario no dispone del conocimiento previo de la solución por lo que intentara agrupar según las relaciones deducidas de los parámetros en tantos grupos como se le indique. Por ultimo Aprendizaje por Refuerzo en el que al sistema se le indica las posibilidades de acción que tiene, la situación inicial, qué sería un fallo o un acierto y una recompensa por cada acierto, de esta forma lo que tiene que buscar el sistema es maximizar la recompensa y este aprenderá de cada error cometido.

Para el caso de la producción de rodamientos se va a disponer de la producción anual con todos los datos de registro de todos los sensores y con el diámetro final obtenido por lo que resulta un claro caso de aplicación de Aprendizaje Supervisado. Dentro de este también se puede dividir en dos subcategorías en este caso según el tipo de soluciones. Por ser el diámetro una solución numérica y no categórica a la producción de rodamientos se abordará como un problema de regresión.

Una de las mejoras más importantes que introduce el uso de este tipo de tecnologías en la industria es la posibilidad del Zero-Defects o cero defectos. Un ejemplo es el proyecto STREAM-OD (1) por el que se pretende la reducción de tiempo, costos y variabilidad de producto mediante la inclusión de sistemas de aporte de soluciones inteligentes basadas en el pronóstico. La predicción permite una mayor eficiencia en la preparación de protocolos de desecho o acondicionamiento de producto así como el ajuste de las variables de producción como método de obtención del resultado deseado. Este proyecto espera conseguir un ahorro de costos de producción del 15% y una reducción de desechos del 10%.

Otro proyecto innovador que también cuenta con el uso de esta tecnología es el Z-Fact0r (2) en el que de un modo parecido busca reducir el número de errores mediante el ajuste de las variables de producción gracias a la rapidez y la capacidad de predicción de Machine Learning. Como rasgo diferenciador de este proyecto con el anterior cabe destacar el quinto objetivo en el que aprovecha la recolección masiva de datos para la extracción de conclusiones e introduciendo técnicas de Business Intelligence permite tomar decisiones de gestión de estrategias. Mediante la construcción de un Datawarehouse se consigue aprovechar toda la recolección de datos de la empresa para poder extraer muchas más conclusiones y poner en valor la información que se dispone gracias a técnicas Business Intelligence.

Todas estas técnicas de análisis de Big Data mejoran conforme más información se dispone de los procesos productivos por lo que el futuro de la industria gira en torno a la aplicación generalizada de este tipo de técnicas sumando la conectividad de las Industrias 4.0 que permiten la formación de alianzas de compartición de información en tiempo real con el objetivo de poder mejorar la toma de decisiones y la competencia como ya se está llevando a cabo en el proyecto BOOST 4.0 (3) con la industria de faros. También se está encaminando el futuro a una mayor accesibilidad de este tipo de técnicas ya que la implementación de estas requieren en muchos casos una especialización y una inversión económica que las PYMES no son capaces de asumir. Con el objetivo de nutrir de estos beneficios a las PYMES surge el proyecto TOREADOR (4) en el

que mediante la estandarización pretende que los costes y la dificultad de aplicación disminuyan generalizando el uso del Big Data.

2. Ajuste modelo simple

Para comenzar se va a estudiar el modelo más simple en el que los diámetros y la holgura se obtendrán con las siguientes fórmulas:

$$A1_6_D1=A1_5_D1.*(1+11.2e-6*(20-A1_6_Tp))+A1_6_I$$

$$A2_3_d1=A2_2_d1.*(1+11.2e-6*(20-A2_3_Tp))+A2_3_EX$$

$$A3_3_d1=A3_2_d1.*(1+11.2e-6*(20-A3_3_Tp))+A3_3_EX$$

$$HOLGURA=(A3_3_d1-100)*4+(A2_3_d1-100)*4-(A1_6_D1-100)*8+(rand(n,1)-0.5)*4/1000+0.060$$

Para la fórmula de la holgura la variabilidad introducida será rectangular para el caso en el que las variables se obtengan con variabilidades rectangulares, y normal cuando las variables tengan también distribución normal.

2.1. Variabilidad rectangular

Se van a clasificar las variables según el siguiente criterio:

- Factores que no influyen
 - o Descubiertos de forma experimental (las gráficas frente a 'n' no muestran tendencia)
 - o Descubiertos de forma analítica (la eliminación de dichas variables para el cálculo del método matemático no influye en su capacidad para predecir el diámetro)
- Factores que influyen
- Factores relacionados (como es el caso de las temperaturas en el que la eliminación de una de ellas puede no afectar al resultado final ya que el modelo lo puede interpretar)

De esta forma se van a ir desechando aquellas variables que no sean útiles para el cálculo.

2.1.1. Aro

Primeramente se va a comprobar si las variables son influyentes experimentalmente, es decir, si muestran una tendencia al representarse gráficamente frente a 'n'.

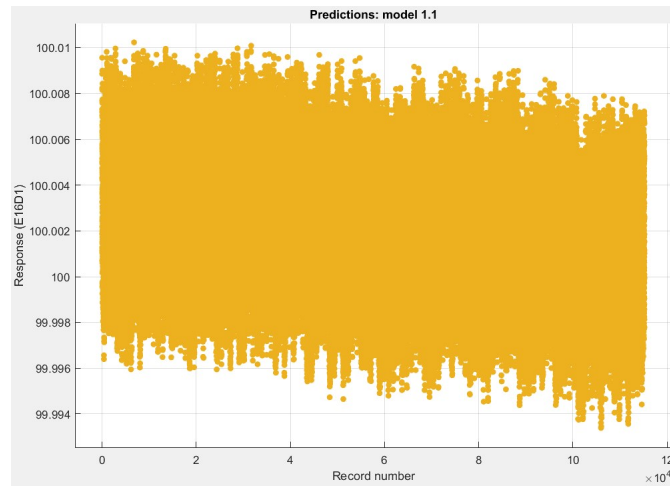


Figura 4. Diámetro frente a n

La muestra de datos que se ha escogido para el aprendizaje del método matemático es medio año con el objetivo de poder tener la otra mitad de datos para poder ejecutar el modelo si se quisiera posteriormente.

La relación presente en la gráfica muestra como conforme se van acercando los meses de verano los diámetros van disminuyendo su tamaño. Esto puede deberse a una posible dependencia con la temperatura.

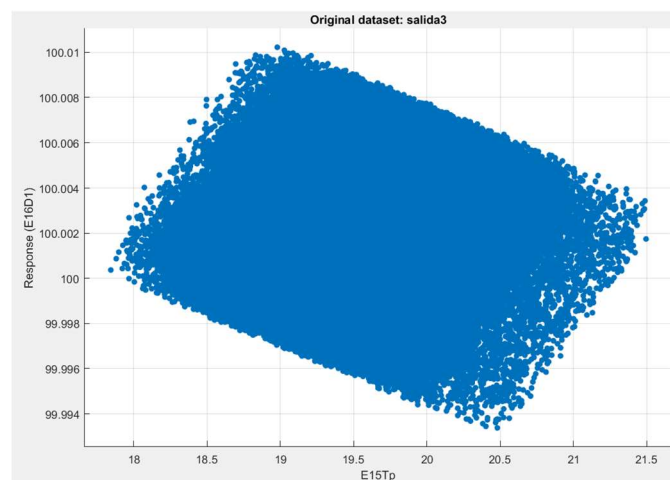


Figura 5. Diámetro frente a T

La temperatura ambiente (A1.6.Ta) al igual que la de la taladrina (S1.6.Tt), el eje x (S1.6.Tx) y la temperatura de la pieza (E1.5.Tp) se comportan de manera similar donde el aumento de dichas variables suponen una disminución en el valor del diámetro. Esto explica la mayor ocurrencia de diámetros más pequeños en los meses de verano.

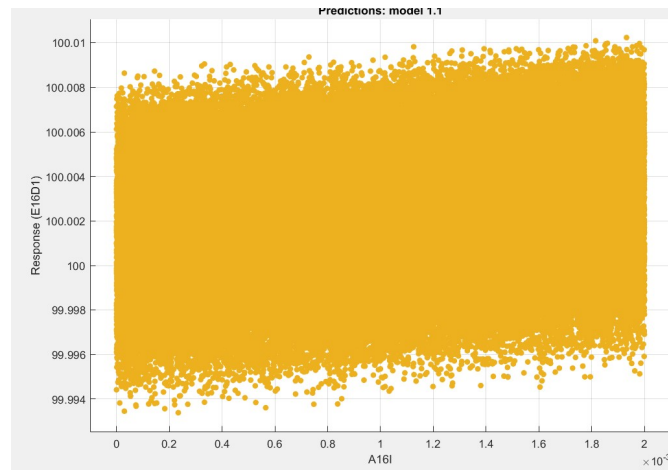


Figura 6. Diámetro frente a incertidumbre

La incertidumbre del palpador (A1.6.I) aunque presenta una tendencia uniforme sí que tiende a aumentar el diámetro final conforme se aumenta la variable.

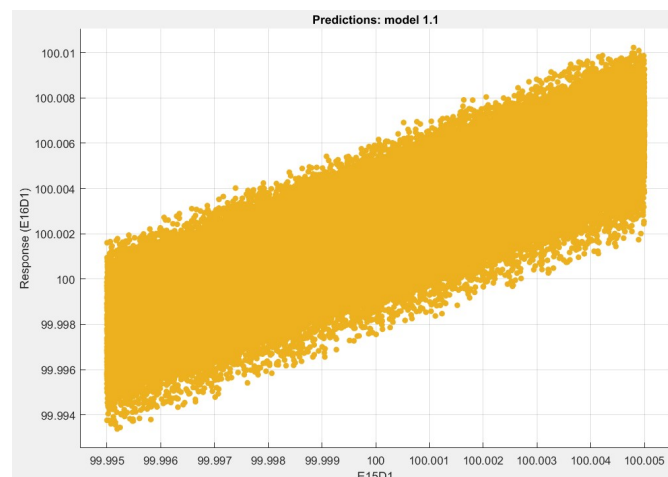


Figura 7. Diámetro frente a E15D1

E1.5.D1 si presenta una tendencia superior que la del palpador representada claramente en la gráfica

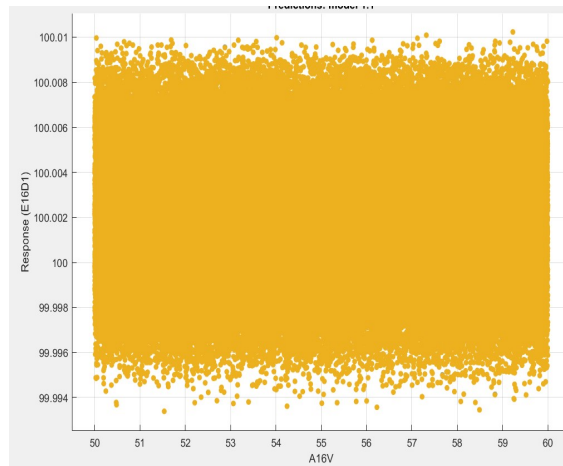


Figura 8. Diámetro frente a vibraciones

Para finalizar el resto de variables presenta una ocurrencia muy similar a la vibración aquí representada en la que se comprueba cómo no son influyentes.

1.1 ☆ Linear Regression	RMSE: 0.00
Last change: Linear	11/11 features
1.2 ☆ Tree	RMSE: 0.00
Last change: Complex Tree	11/11 features
1.3 ☆ Tree	RMSE: 0.00
Last change: Medium Tree	11/11 features
1.4 ☆ Tree	RMSE: 0.00
Last change: Simple Tree	11/11 features

Figura 9. Errores de los modelos

Para la simulación se ha comparado los cuatro modelos más rápidos de entrenar dando todos ellos un error cuadrático de cero por lo que se deberá tener en cuenta otros factores a la hora de decidir qué modelo es el más adecuado.

2.1.1.1. Linear regression

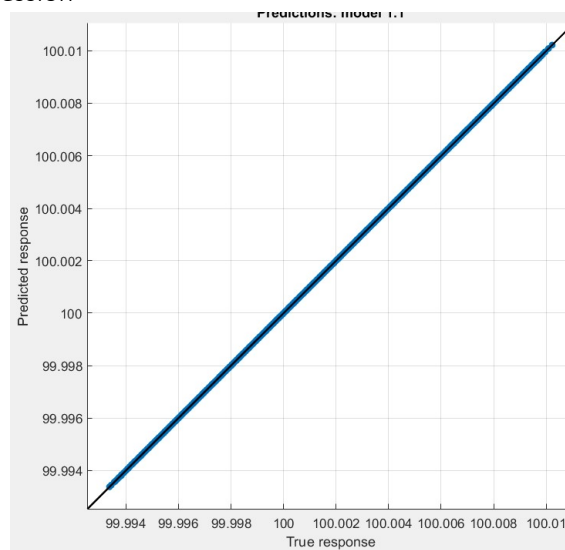


Figura 10. Respuesta predicha frente a real

Como se puede observar, y como fruto de que el diámetro final lo hemos obtenido de una fórmula matemática, la predicción corresponde perfectamente con la respuesta verdadera.

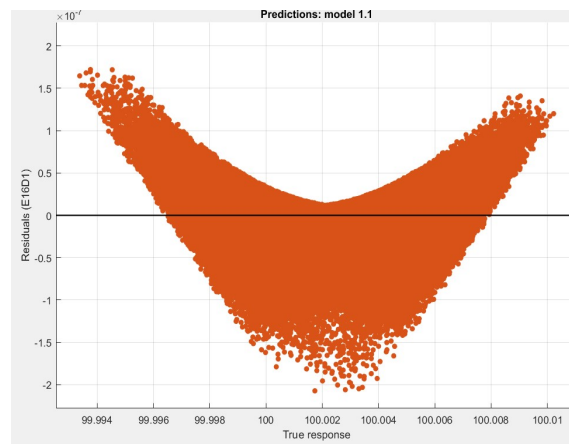


Figura 11. Ocurrencia de valores residuales

Para intentar decidir el modelo a escoger compararemos la representación de los valores residuales que deberán aparecer de manera uniforme alrededor de 0 en toda la horizontal como forma ideal.

En este modelo se observa cómo se queda lejos de la uniformidad aunque los valores son del orden de 10^{-7} .

2.1.1.2. Complex tree

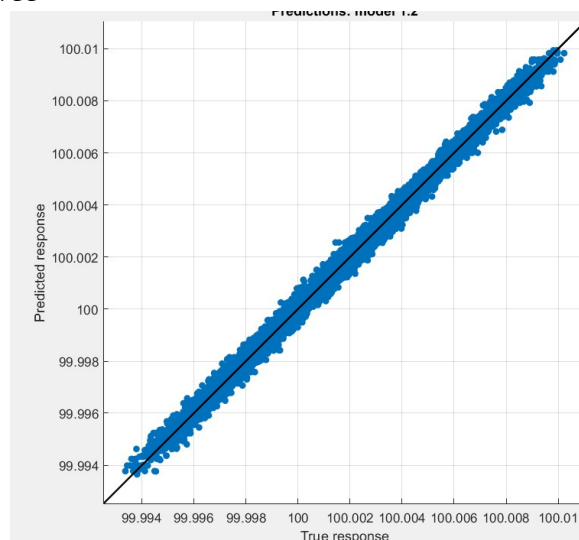


Figura 12. Respuesta predicha frente a real (2)

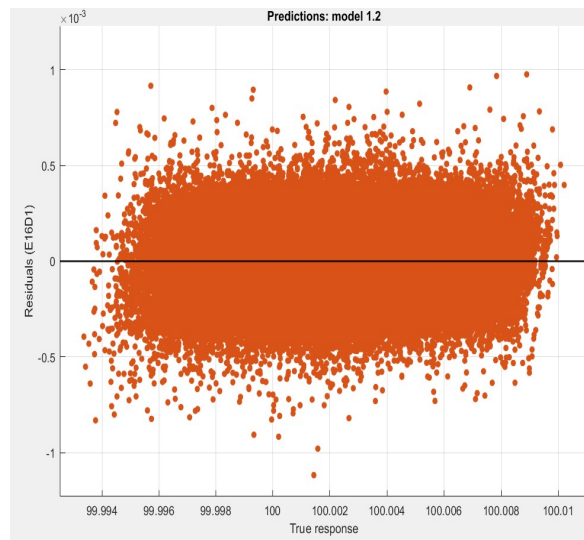


Figura 13. Ocurrencia de valores residuales (2)

Al ser este un método más simple que el anterior la respuesta predicha no es tan perfecta pero sigue siendo una muy buena aproximación con un RMSE de cero. En este modelo la representación de los valores residuales sí que sigue una distribución uniforme con respecto a cero lo cual es beneficioso. Estos valores residuales sin embargo son de un orden menor a 10^{-3} lo que puede ser un error aceptable teniendo en cuenta que se ha partido de datos con tres decimales.

Por tanto este modelo presenta una complejidad de computación mejor pero presenta unos valores residuales mayores.

2.1.1.3. Simple y medium tree

Con estos modelos, aunque el error cuadrático es de cero, los valores residuales también alcanzan el orden de 10^{-3} lo que puede suponer un error de predicción no asumible para el proceso de fabricación.

2.1.1.4. Cribado de factores

A continuación se va a comprobar la influencia analítica de las variables. Primero se eliminan del modelo aquellas variables que se han clasificado como no influyentes de forma experimental (vibración, ruido, error en la consigna del palpador y el consumo de potencia), ya que no afectaran al cálculo del modelo. De esta forma solo quedarán las temperaturas (variables relacionadas) y la incertidumbre como variables.

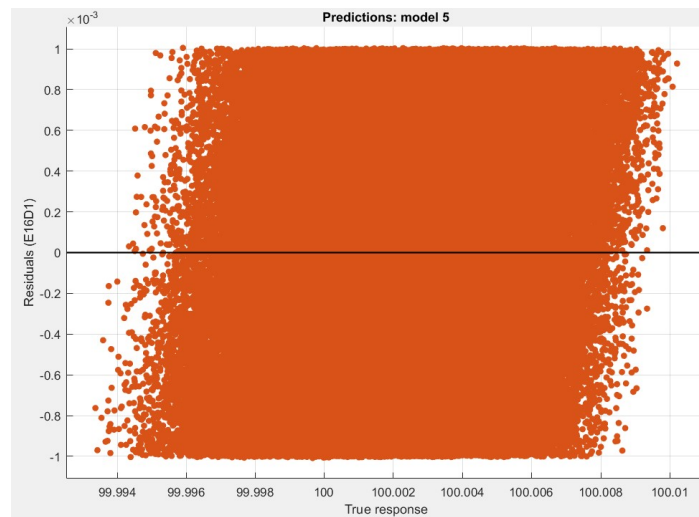


Figura 14. Ocurrencia de valores residuales sin considerar la incertidumbre

Al eliminar como variable la incertidumbre el modelo presenta un empeoramiento de los resultados ya que pasa de un error de cálculo del orden de 10^{-7} a menor de 10^{-3} por lo que esta es una variable influyente. No obstante para este caso, con el método *Linear Regression* se podría eliminar ya que el error cometido seguiría por debajo de la milésima.

2.1.1.5. Introducción de ruido

Se va a comprobar la robustez del método mediante la adición de un ruido independiente. Para ver como el ruido afecta a la capacidad del método para predecir el modelo se va a llevar a cabo una gráfica representando el valor de porcentaje introducido frente al error cuadrático cometido en la predicción (RMSE). La cantidad de ruido introducido se obtiene de la relación entre la desviación típica del ruido introducido frente la del diámetro. Para la realización de estas gráficas se va a tomar las RMSE producidas con los dos mejores modelos antes observados (*Linear Regression* y *Tree*) para comprobar su evolución.

s1	s2	s1/s2 (%)	Tree	Linear Regression
			RMS	RMS
0,0058	0,0013	22,52	6,21E-04	5,27E-04
0,0114	0,0013	11,40	1,20E-03	1,04E-03
0,0084	0,0013	15,48	9,06E-04	7,75E-04
0,0291	0,0013	4,47	3,09E-03	2,65E-03

Tabla 4. RMSE frente a porcentaje de ruido

Siendo S1 la desviación del ruido introducido y S2 la del diámetro sin ruido.

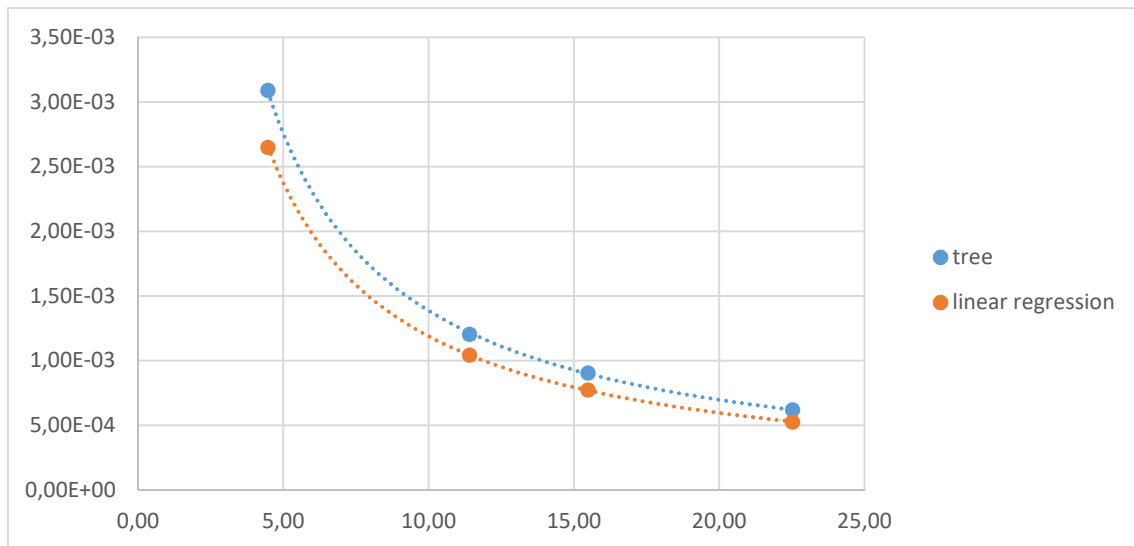


Figura 15. Representación RMSE frente a porcentaje de ruido

Se puede observar dónde se produce el cambio en el que la RMSE cambia de orden de magnitud al de las micras para ambos métodos matemáticos.

2.1.2. Cono A y cono B

En este caso ambos se comportarán de igual manera ya que se han generado los datos de igual manera por lo que se tratarán en este mismo apartado.

Se realiza el mismo estudio de influencia de las variables que en el apartado anterior. En este caso la todas las variables se comportan de forma similar que lo hacían en el aro.

1.1	☆ Linear Re...	RMSE: 0.00
Last change: Linear		11/11 features
1.2	☆ Tree	RMSE: 0.00
Last change: Compl...		11/11 features
1.3	☆ Tree	RMSE: 0.00
Last change: Mediu...		11/11 features
1.4	☆ Tree	RMSE: 0.00
Last change: Simple...		11/11 features

Figura 16. Errores de los modelos (2)

Como se observa los errores, al igual que en el caso anterior, son igual a 0 por lo que en cuanto a la selección del mejor método se seguirá el mismo criterio que en el caso anterior. Los valores residuales para el caso *Linear regression* vuelven a ser del orden de 10^{-7} frente a los modelos *tree* en los que son del orden de 10^{-4} .

2.1.2.1. Linear regression

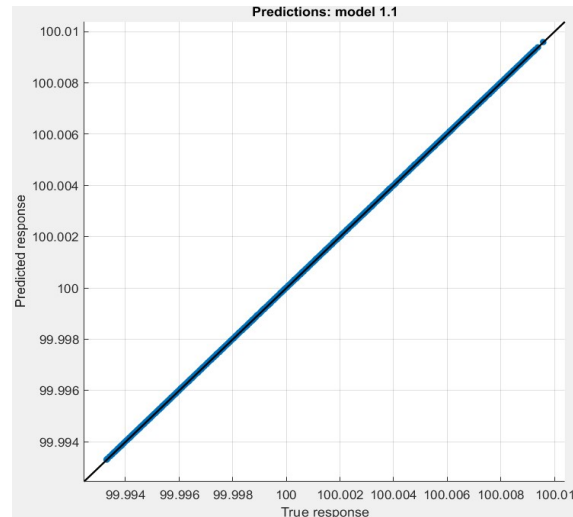


Figura 17. Respuesta predicha frente a real (3)

Al igual que en el aro este modelo es el que mejor consigue realizar la predicción de los valores con unos valores residuales muy bajos con respecto a las medidas tomadas aunque no tengan la mejor distribución deseada.

El análisis de los conos será igual para todos los casos al análisis del aro ya que ambos se obtienen de la misma manera y con fórmulas equivalentes por lo que se procederá a analizar solo los resultados obtenidos del aro ya que serán extrapolables a los conos.

2.1.3. Holgura

El cálculo de la holgura se hace a través de los tres diámetros de los que está compuesto el rodamiento. No obstante para el estudio experimental y analítico se van a tomar como variables de entrada todas las variables utilizadas previamente para el estudio de los diámetros de los aros. De esta forma se asemejará el estudio al estudio real ya que la holgura total dependerá de los parámetros de fabricación de los aros.

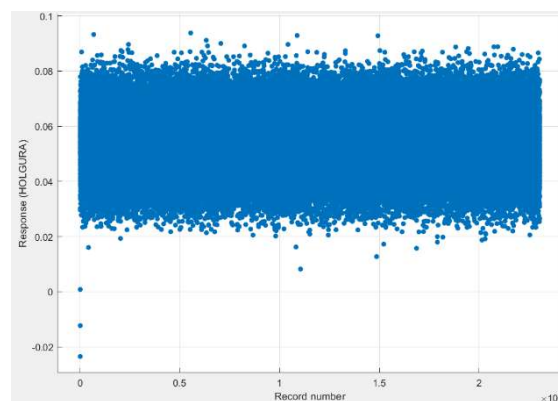


Figura 18. Holgura frente a n

Como se puede observar en la representación de la holgura frente a "n" no se observa tendencia como si ocurría en el aro con la presencia del verano.

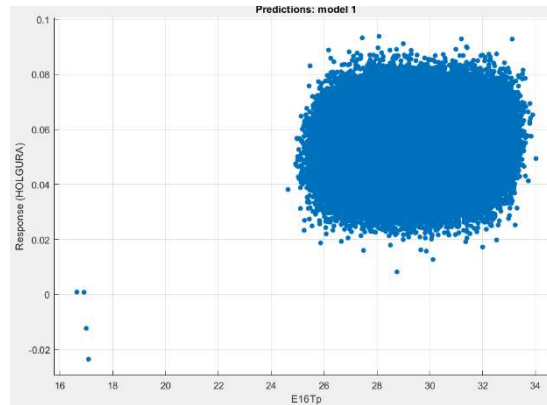


Figura 19. Holgura frente a T

Esto queda explicado en la representación de la temperatura de la pieza frente a la holgura en la que queda reflejada esa nula relación. De igual forma ocurre con el resto de las temperaturas.

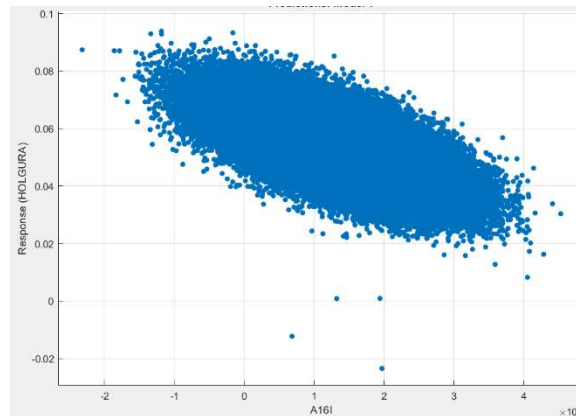


Figura 20. Holgura frente a incertidumbre

La incertidumbre presenta una clara dependencia con el diámetro donde el aumento de esta provoca la aparición de diámetros más pequeños al igual que la variable E15D1.

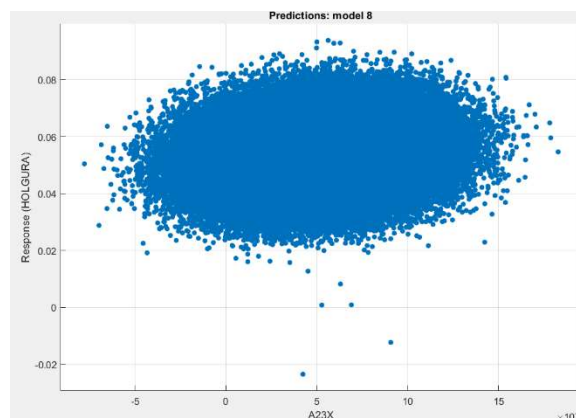


Figura 21. Holgura frente a consigna palpador

Esta variable que ha sido irrelevante en el aro tiene una ligera dependencia experimental por lo que será necesario evaluarla con más detenimiento en la fase analítica para comprobar su dependencia real y cómo de significativa es.

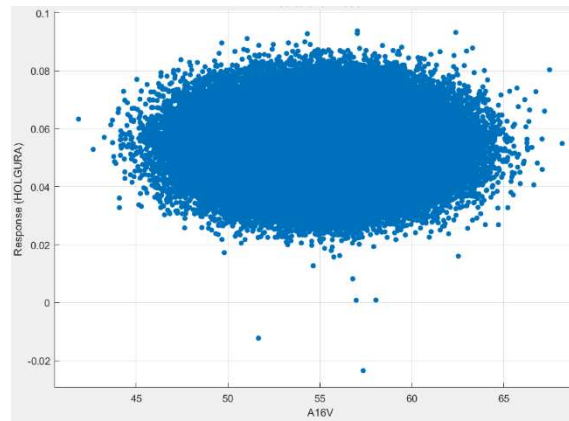


Figura 22. Holgura frente a vibración

Para finalizar el resto de variables no muestran una influencia experimental como se muestra en la representación de las vibraciones.

2.1.3.1. Cribado de factores

Para este caso y al tener muchas más variables que en el caso del aro se va a proceder a ir analizando los errores cuadráticos (RMSE) probando cada variable relevante en la fase experimental.

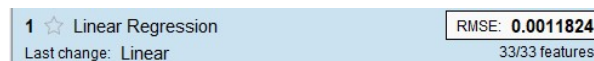


Figura 23. RMSE considerando todas las variables

Para comenzar se toma como referencia el método en el que se consideran todas variables para poder comparar posteriormente.



Figura 24. RMSE considerando incertidumbre del palpador y los tres diámetros

Considerando solo las variables que en el análisis experimental eran indudablemente influyentes se produce un aumento de tres veces el error cometido anteriormente por lo que será necesario tener en cuenta más variables.



Figura 25. RMSE considerando incertidumbre, diámetros y una temperatura de cada pieza

Incluyendo una temperatura de cada pieza se ha conseguido disminuir considerablemente el error de predicción cometido pero se va a eliminar una temperatura con el objetivo de cerciorarse de que es necesario una temperatura de cada pieza.



Figura 26. RMSE considerando incertidumbre, diámetros y dos temperaturas de dos piezas

Efectivamente se demuestra que es necesaria la consideración de una temperatura de cada pieza para realizar la predicción de forma más exacta.

6 ☆ Linear Regression RMSE: 0.0011827
Last change: Added 2 features 10/33 features

Figura 27. RMSE considerando incertidumbre, diámetros, una temperatura de cada pieza y las tres X

Para finalizar se comprueba que aunque la dependencia experimental de la consigna del palpador era muy ligera si resulta relevante analíticamente para rebajar el error.

2.1.3.2. Introducción de ruido

Al igual que con el aro se va a analizar el método con la adición de un ruido:

s1	s2	s1/s2 (%)	tree	linear regression
			RMS	RMS
0,0058	0,0088	152,42	7,02E-03	4,40E-03
0,0029	0,0088	304,84	5,49E-03	2,45E-03
0,0019	0,0088	457,26	5,17E-03	1,88E-03
0,0014	0,0088	609,68	5,03E-03	1,59E-03
0,0012	0,0088	762,10	4,970E-03	1,466E-03
0,0006	0,0088	1524,20	4,884E-03	1,277E-03

Tabla 5. RMSE frente a porcentaje de ruido (2)

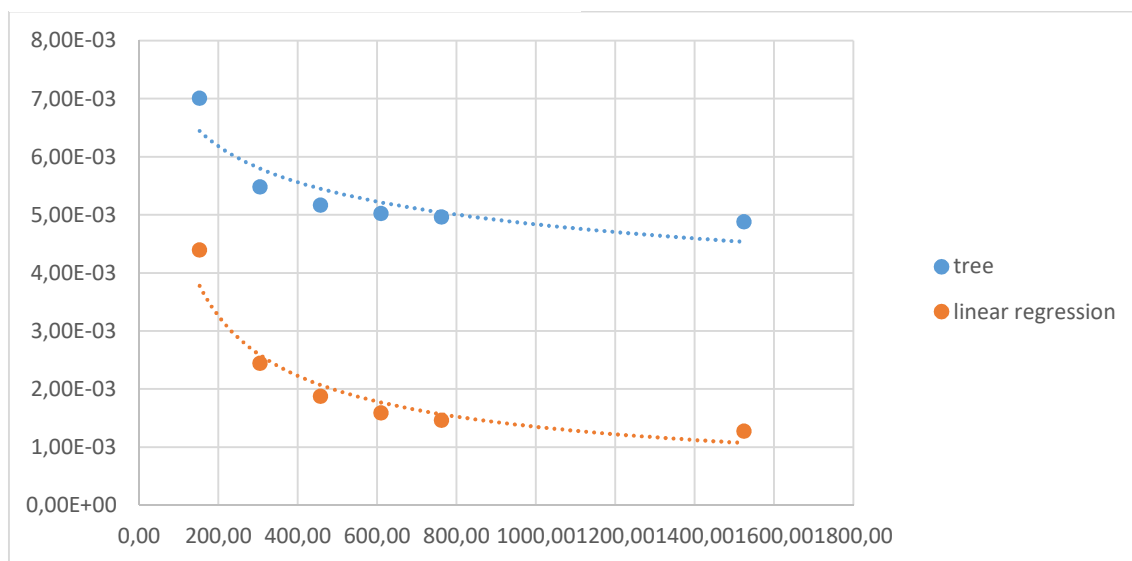


Figura 28. Representación RMSE frente a porcentaje de ruido (2)

Para este caso si se observa más diferencias entre un método y en otro ya que la diferencia entre los dos en la fase en la que se estabiliza el de casi medio orden de magnitud.

2.2. Variabilidad normal

En los casos anteriores la evolución de los datos generados seguía una distribución rectangular, para este caso se va a comprobar cómo afectará el cambio a una distribución normal.

2.2.1. Aro

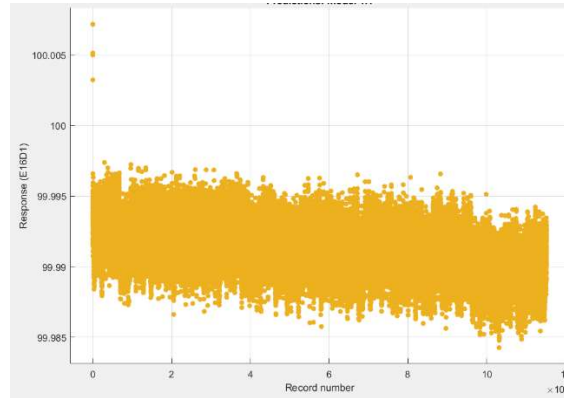


Figura 29. Diámetro frente a n (2)

Como se puede observar en la distribución de datos a lo largo de medio año también existe una tendencia en la que en los meses de mas calor los diámetros son más pequeños. Propio de la distribución normal es la aparición, en una pequeña proporción, de datos más extremos como en este caso ocurre concentrados en los primeros días.

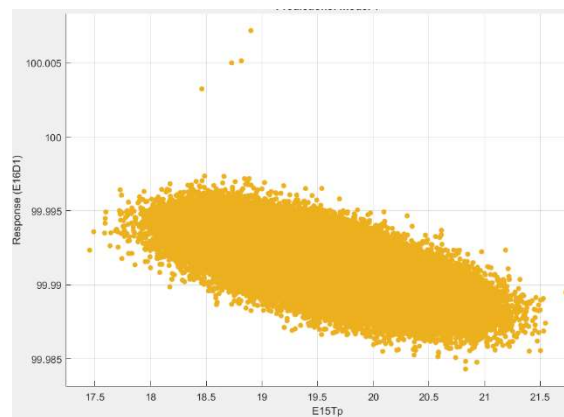


Figura 30. Diámetro frente a T (2)

Esta influencia de la temperatura en los meses se ve reflejada cuando se representan las diferentes temperaturas frente al diámetro en la que se aprecia esta tendencia en la que a mayor temperatura la tendencia del diámetro es a disminuir.

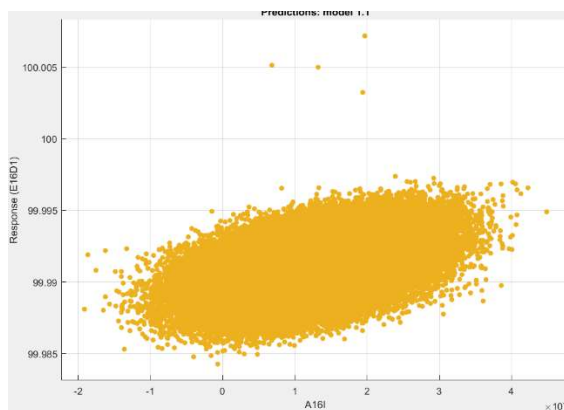


Figura 31. Diámetro frente a incertidumbre (2)

Para este caso la incertidumbre, que anteriormente tenía una ligera influencia, si que presenta una clara tendencia en la que a medida que esta aumenta lo hace también el diametro.

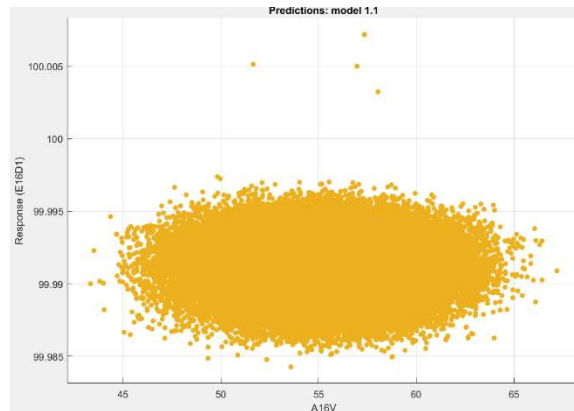


Figura 32. Diámetro frente a vibraciones (2)

El resto de variables no influyen en el diametro final de la pieza al igual que el representado como es la vibración.

2.2.1.1. Cribado de factores

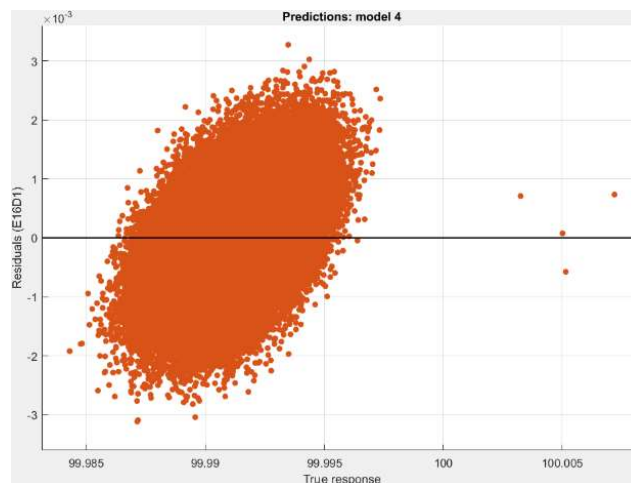


Figura 33. Ocurrencia de valores residuales sin considerar la incertidumbre (2)

Como se ha observado la dependencia de las temperaturas hace que solo teniendo en cuenta una de ellas para la predicción se cometa el mismo error que considerando todas ellas. En el caso de la incertidumbre del palpador se demuestra su influencia analítica más allá de la experimental ya que su eliminación eleva el error cometido a un orden de 10^{-3} .

2.2.1.2. Introducción de ruido

Al igual que en el modelo con variabilidad rectangular se analizará la inclusión del ruido de la misma manera:

s1	s2	s1/s2 (%)	Tree	linear regression
			RMS	RMS
0,0058	0,0020	34,64	7,50E-03	5,77E-03
0,0029	0,0020	69,28	3,75E-03	2,88E-03
0,0019	0,0020	103,92	2,52E-03	1,93E-03
0,0014	0,0020	138,56	1,89E-03	1,45E-03
0,0012	0,0020	173,21	1,519E-03	1,159E-03
0,0006	0,0020	346,41	7,641E-04	5,776E-04
0,0004	0,0020	519,62	5,186E-04	3,858E-04
0,0003	0,0020	692,82	3,961E-04	2,879E-04
0,0002	0,0020	866,03	3,251E-04	2,302E-04

Tabla 6. RMSE frente a porcentaje de ruido (3)

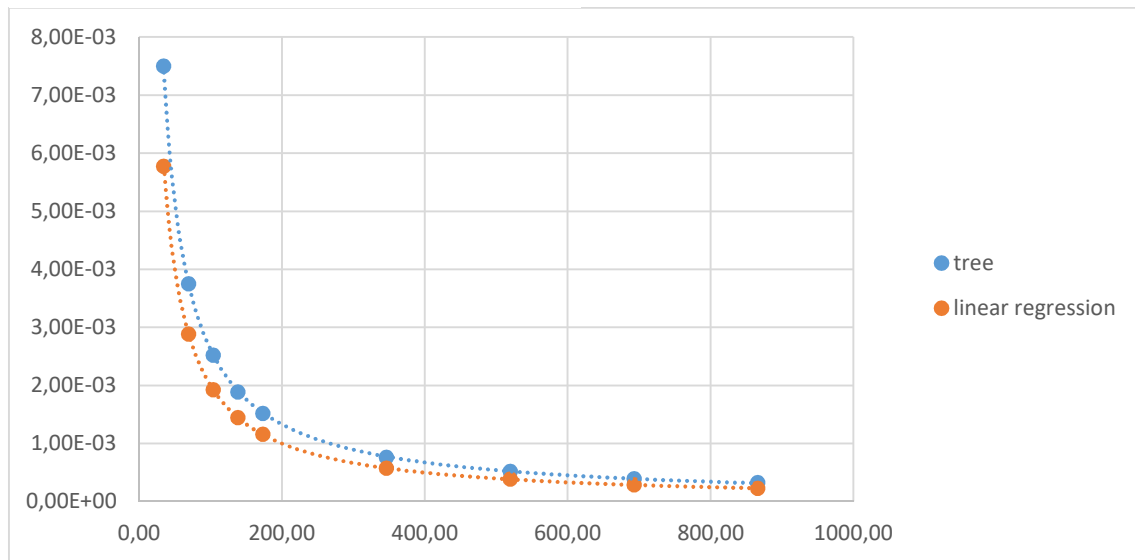


Figura 34. Representación RMSE frente a porcentaje de ruido (3)

Para este caso también se ha realizado un análisis de todos los métodos matemáticos disponibles en la herramienta *Linear Regression* para cerciorarse de que los dos métodos que se están escogiendo son los mejores para predecir este proceso de fabricación. De esta forma se han obtenido todas las RMSE de todos los métodos para dos porcentajes de ruido.

Para $s_1/s_2=34.64$:

1.1 ☆ Linear Regression	RMSE: 0.0057725
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0057737
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.0057725
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0057732
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0075027
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.0066863
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0061336
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.0057742
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0057781
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0057904
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0060315
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0058232
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0057776
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2387
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0058606
Last change: Bagged Trees	11/11 features

Figura 35. RMSE para todos los modelos de Regression Learner

Para $s_1/s_2=173.21$:

1.1 ☆ Linear Regression	RMSE: 0.001159
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0011595
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.001159
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0011591
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0015187
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.0013504
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0012411
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.0011607
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0011634
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0011648
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0015022
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0011751
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0011637
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2387
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0011872
Last change: Bagged Trees	11/11 features

Figura 36. RMSE para todos los modelos de Regression Learner (2)

Donde se observa que para distintos niveles de ruido los métodos que mejor describen dicho proceso son los de regresión y los de árbol.

2.2.2. Holgura

Como en el modelo para una variabilidad rectangular también se va a analizar la holgura. Inicialmente se va a hacer el estudio experimental de las variables.

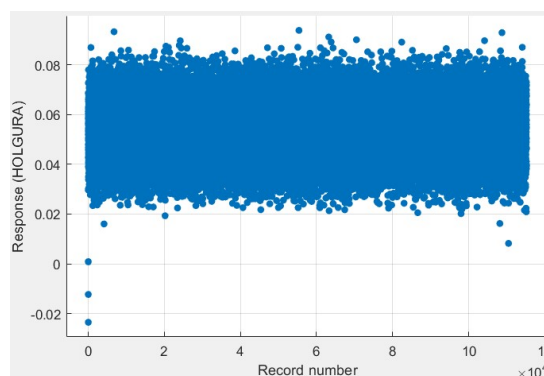


Figura 37. Holgura frente a n (2)

Al igual que en el otro caso que se ha analizado la holgura no se aprecia una dependencia de dicha holgura con el paso del tiempo ' n '.

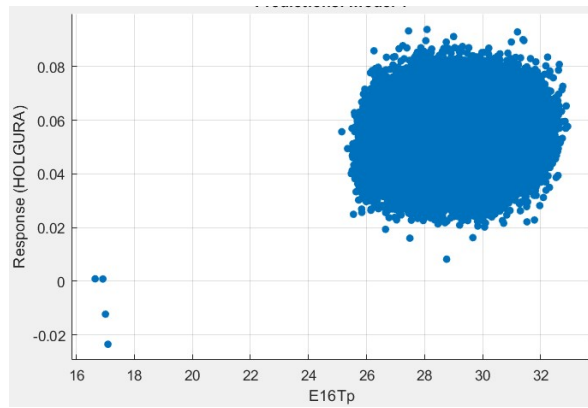


Figura 38. Holgura frente a T (2)

La nula relación antes descrita se puede explicar con la aparente nula dependencia de la temperatura en el resultado de la holgura, no obstante ya se ha demostrado en el caso anterior como si guardaba una relación analítica.

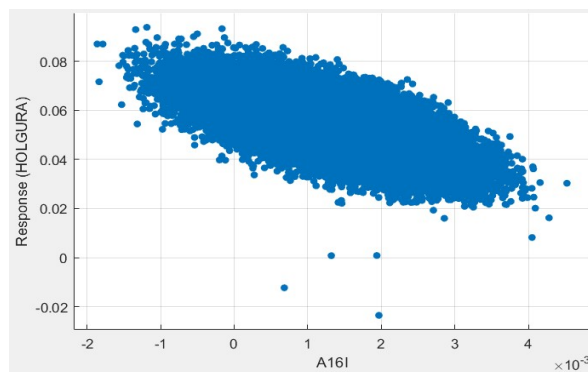


Figura 39. Holgura frente a incertidumbre (2)

Tanto la incertidumbre como la variable E15D1 guardan una clara relación con el diámetro donde el aumento de estas provoca una disminución en el diámetro.

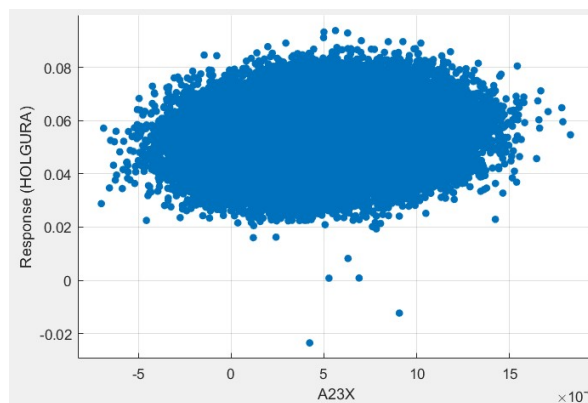


Figura 40. Holgura frente a consigna del palpador (2)

Al igual que en el caso del análisis de la holgura para variabilidad rectangular la consigna del palpador guarda una ligera relación con el diámetro.

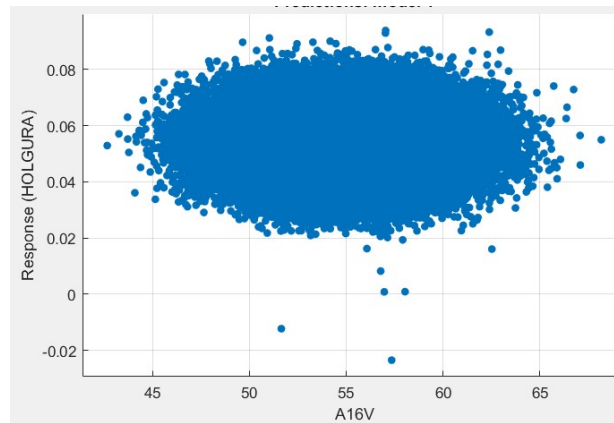


Figura 41. Holgura frente a vibración (2)

Para finalizar el resto de variables no guardan una relación experimental como se puede apreciar con la representación de la vibración.

2.2.2.1. Cribado de factores

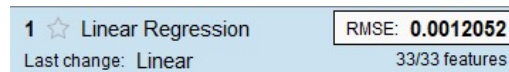


Figura 42. RMSE considerando todas las variables (2)

Se tomará el valor de RMSE considerando todas las variables como valor de referencia para comparar el resto de simulaciones.

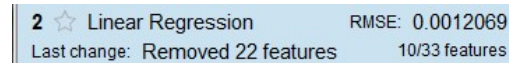


Figura 43. RMSE considerando incertidumbre, diámetros, una temperatura de cada pieza y las tres X (2)

Teniendo en cuenta los resultados del análisis experimental, los cuales eran similares a los de la holgura con variabilidad rectangular, y tomando lo aprendido en ese caso se comprueba que el escoger las mismas variables también es válido para este caso.

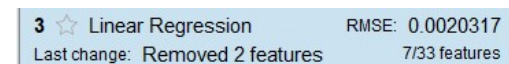


Figura 44. RMSE considerando incertidumbre, diámetros y dos temperaturas de dos piezas (2)

No obstante se ha hecho otra simulación para comprobar que esa es la mejor selección de variables para este caso.

3. Ajuste modelo complejo

A partir del modelo simple se han desarrollado más las fórmulas de donde se extraían los diámetros y la holgura dando lugar a las siguientes:

$$A1_6_D1 = ((\text{mean}(A1_5_D1) * (1 + 11.2e-6 * (20 - (A1_5_Tp * 0.2 + A1_6_Tp * 0.8)))) * (1 + 11.2e-6 * ((A1_6_Tt * 0.2 + A1_6_Tx * 0.8) - 20)) + A1_6_I);$$

$$A2_3_d1 = ((\text{mean}(A2_2_d1) * (1 + 11.2e-6 * (20 - (A2_2_Tp * 0.2 + A2_3_Tp * 0.8)))) * (1 + 11.2e-6 * ((A2_3_Tt * 0.2 + A2_3_Tx * 0.8) - 20)) + A2_3_EX);$$

$$A3_3_d1 = ((\text{mean}(A3_2_d1) * (1 + 11.2e-6 * (20 - (A3_2_Tp * 0.2 + A3_3_Tp * 0.8)))) * (1 + 11.2e-6 * ((A3_3_Tt * 0.2 + A3_3_Tx * 0.8) - 20)) + A3_3_EX);$$

$$HOLGURA = (A3_3_d1 - 100) * 4 + (A2_3_d1 - 100) * 4 - (A1_6_D1 - 100) * 8 + (\text{rand}(n, 1) - 0.5) * 4 / 1000 + 0.060;$$

Para la fórmula de la holgura la variabilidad introducida será rectangular para el caso en el que las variables se obtengan con variabilidades rectangulares, y normal cuando las variables tengan también distribución normal.

Al igual que en el modelo simple se estudiará para este modelo cómo se comporta ante dos tipos de distribución en la ocurrencia de las variables: rectangular y normal.

3.1. Variabilidad rectangular

3.1.1. Aro

Al igual que se ha demostrado en el modelo simple la metodología para la obtención tanto de las variables como de los diámetros es similar para el aro y los conos por lo que el análisis que se va a realizar para el aro se podrá extrapolar a los conos.

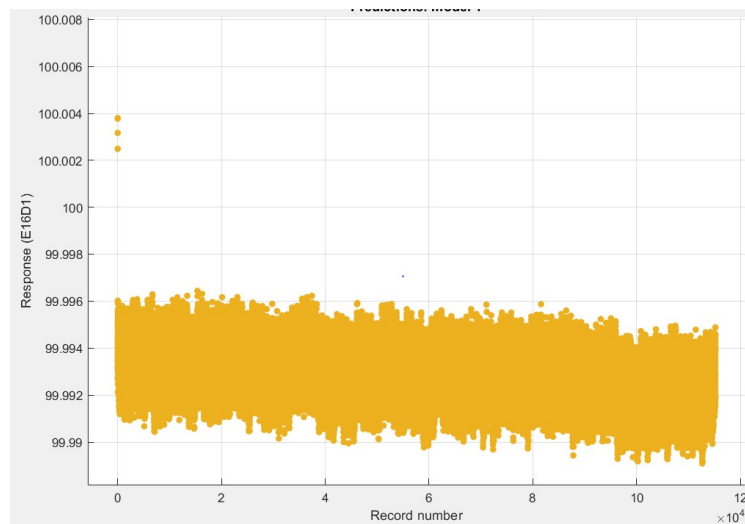


Figura 45. Diámetro frente a n (3)

Se puede apreciar la tónica habitual en la que en los meses de verano los diámetros producidos son menores que el resto del año.

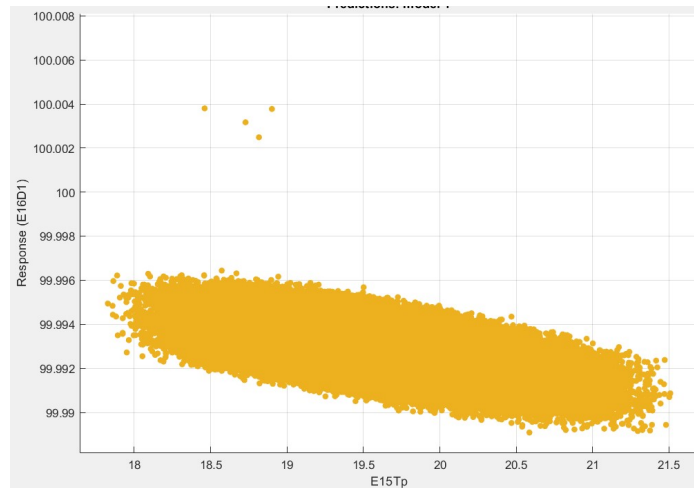


Figura 46. Diámetro frente a T (3)

La relación anteriormente descrita se aprecia en la influencia que la temperatura tiene en el diámetro como se representa en la gráfica anterior.

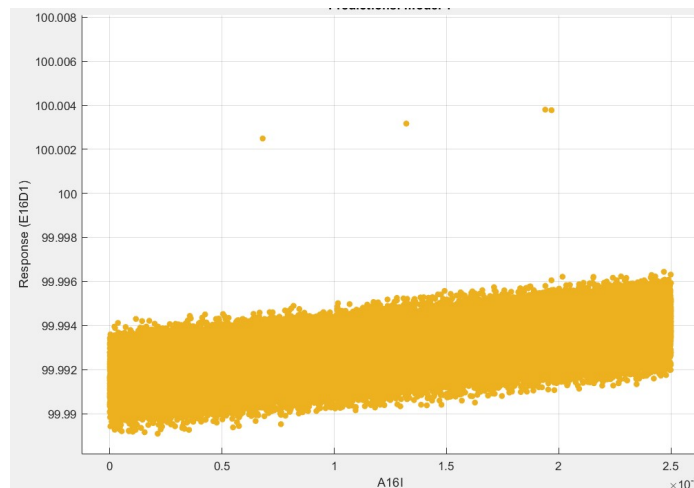


Figura 47. Diámetro frente a incertidumbre (3)

La incertidumbre también guarda influencia con el diámetro al revés que la temperatura en donde un aumento de la incertidumbre provoca unos diámetros mayores.

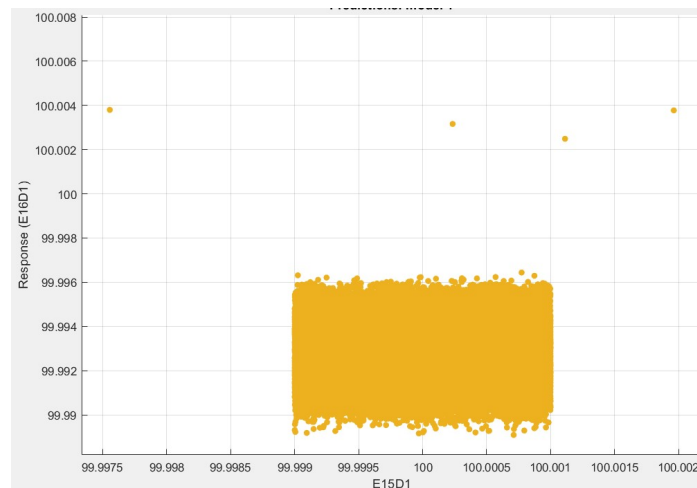


Figura 48. Diámetro frente a E15D1 (2)

Al contrario que lo que venía ocurriendo la variable E15D1 en este caso deja de presentar influencia en el diámetro final. El resto de variables tampoco presentan influencia sobre el resto de variables.

3.1.1.1. Cribado de factores



Figura 49. RMSE considerando todas las variables (3)

Para el cribado de factores se va a partir del valor de la RMSE con la consideración en el cálculo de todas las variables para poder comparar posteriormente la influencia analítica del resto de variables.



Figura 50. RMSE considerando una temperatura y la incertidumbre

Inicialmente se ha simulado la opción más satisfactoria utilizada en el modelo simple en el que con tan solo una temperatura y la incertidumbre se conseguía que el modelo matemático consiguiera predecir los diámetros cometiendo el mismo error que con todas las variables. Esta opción demuestra que, aunque el valor de error cometido sea pequeño, algunas variables que son influyentes no se están teniendo en cuenta.

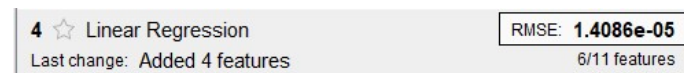


Figura 51. RMSE considerando todas las temperaturas y la incertidumbre

Incluyendo todas las temperaturas como variables a tener en cuenta se logra reducir el error cometido hasta valor de error inicial, por lo que son variables influyentes. No obstante se va a comprobar si todas las temperaturas son influyentes.

5 ☆ Linear Regression	RMSE: 1.4068e-05
Last change:	5/11 features

Figura 52. RMSE considerando la incertidumbre y todas las temperaturas menos T_a

Con la eliminación de T_a de las variables influyentes se logra mantener el mismo error por lo que se deduce que esta variable no es influyente analíticamente. Esto puede deberse a que esta variable no está incluida en las fórmulas para hallar los diámetros.

6 ☆ Linear Regression	RMSE: 7.2735e-05
Last change:	4/11 features

Figura 53. RMSE considerando la incertidumbre y todas las temperaturas menos T_a y T_t

El resto de temperaturas sí que son influyentes analíticamente ya que su eliminación provoca un aumento en el error de predicción.

3.1.1.2. Introducción de ruido

s1	s2	s1/s2	tree	linear regression
			RMS	RMS
0,0058	0,0012	20,78	7,56E-03	5,79E-03
0,0029	0,0012	41,57	3,77E-03	2,88E-03
0,0019	0,0012	62,35	2,52E-03	1,92E-03
0,0014	0,0012	83,14	1,89E-03	1,45E-03
0,0012	0,0012	103,92	1,511E-03	1,154E-03
0,0006	0,0012	207,85	7,618E-04	5,778E-04
0,0003	0,0012	415,69	3,961E-04	2,891E-04

Tabla 7. RMSE frente a porcentaje de ruido (4)

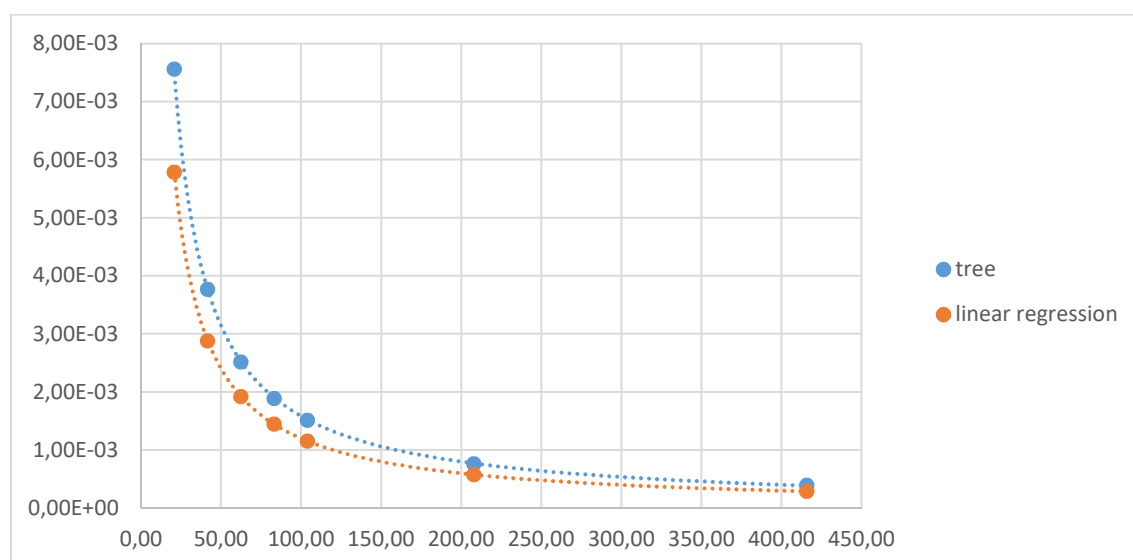


Figura 54. Representación RMSE frente a porcentaje de ruido (4)

Se ha estudiado la influencia de la introducción de un ruido independiente en el cálculo del error para el modelo con un error menor y para el modelo con un tiempo de computación menor. Se

puede apreciar el punto en los dos modelos en el que la RMSE pasa a ser de un orden menor al de las milésimas.

Con $S1/S2=20,78$:

1.1 ☆ Linear Regression	RMSE: 0.0057892
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0057912
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.0057892
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0057894
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0075624
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.006729
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0061531
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.0057914
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0058018
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0058197
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0061004
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0058709
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0057982
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2388
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0058692
Last change: Bagged Trees	11/11 features

Figura 55. RMSE para todos los modelos de Regression Learner (3)

Con $S1/S2=1,511E-03$:

1.1 ☆ Linear Regression	RMSE: 0.0011541
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0011545
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.0011542
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0011543
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0015156
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.0013473
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0012316
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.001157
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0011603
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0011632
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0013014
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0011733
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0011583
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2388
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0011779
Last change: Bagged Trees	11/11 features

Figura 56. RMSE para todos los modelos de Regression Learner (4)

Para confirmar que los modelos matemáticos escogidos para todos los cálculos siguen siendo óptimos se ha realizado una comparativa en dos estados diferentes de ruido con todos los modelos disponibles en la aplicación. Se confirma la elección de *Linear Regression* como modelo más fiable y que el modelo *Tree* a pesar de cometer un error mayor, por tener más ahorro de computación, tampoco comete un error muy alto comparado con el resto.

3.1.2. Holgura

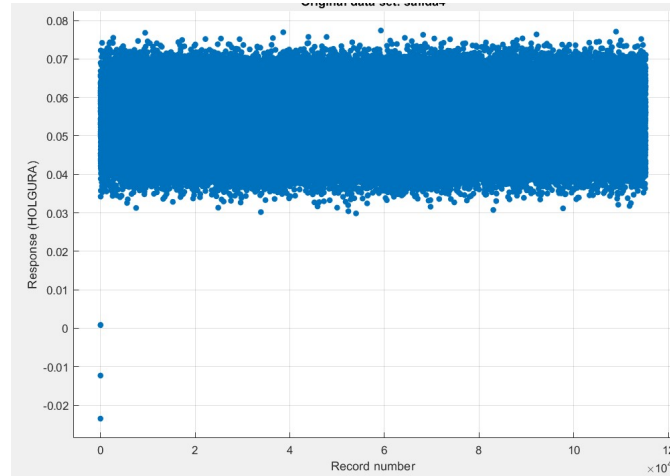


Figura 57. Holgura frente a n (3)

En la representación de la holgura frente a 'n' al igual que en el modelo simple esta no presenta relación con el paso del tiempo.

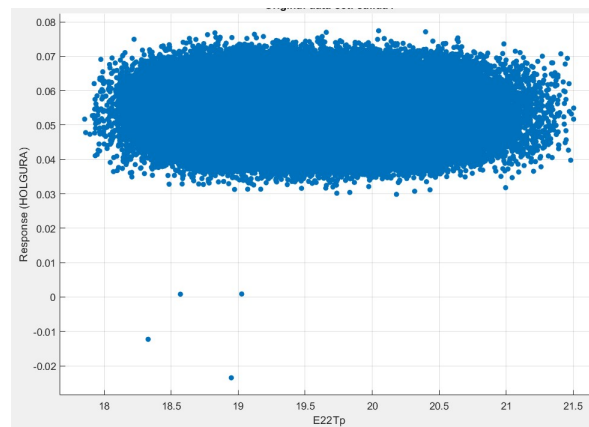


Figura 58. Holgura frente a T (3)

Lo anteriormente comentado se refleja en la nula influencia que experimentalmente se aprecia de la temperatura en el diámetro.

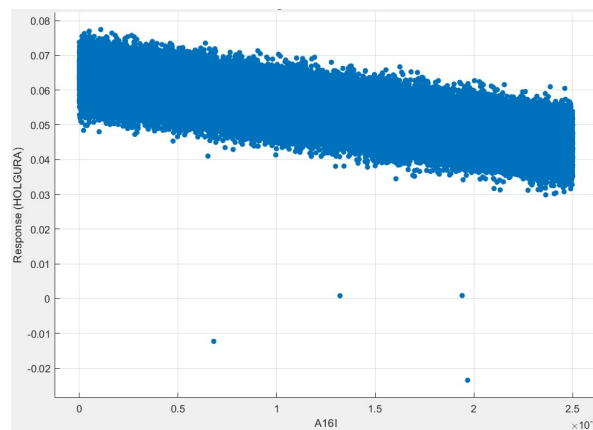


Figura 59. Holgura frente a incertidumbre (3)

Como en todos los casos tratados la influencia de la incertidumbre sigue siendo relevante pero en este caso es al revés en el que el aumento de esta provoca diámetros menores.

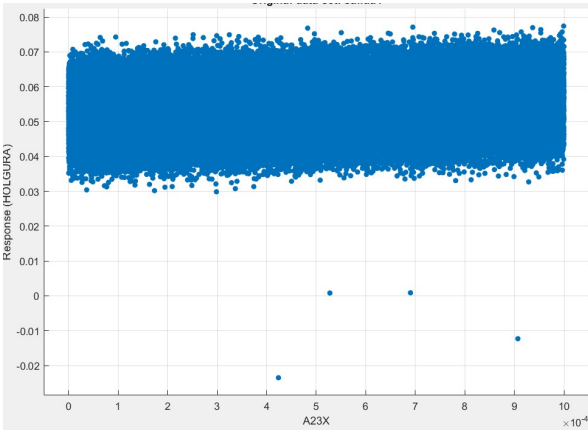


Figura 60. Holgura frente a consigna palpador (3)

La variable consigna en el eje x presenta una ligera dependencia con el diámetro.

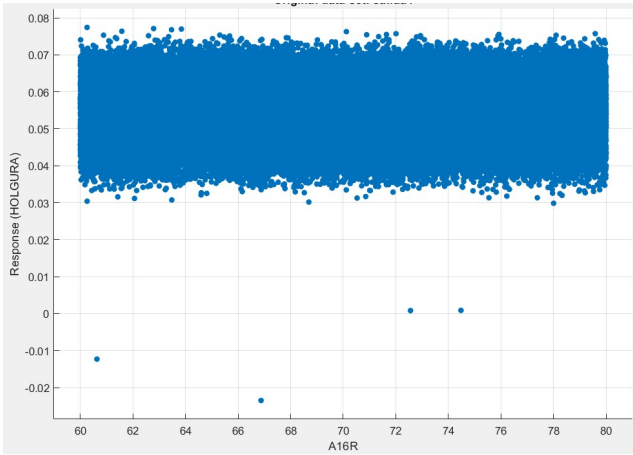


Figura 61. Holgura frente a ruido

El resto de variables no presentan influencia en el estudio experimental.

3.1.2.1. Cribado de factores



Figura 62. RMSE considerando todas las variables (4)

Se inicia el estudio del cribado de los factores en este apartado con el cálculo del valor referencial de la RMSE.



Figura 63. RMSE considerando incertidumbre, diámetros, una temperatura de cada pieza y las tres X (3)

Al igual que para el apartado del aro se aplica la solución adoptada en el modelo simple para la holgura resultando también imprecisa para el modelo actual de estudio.

Figura 64. RMSE considerando incertidumbre, diámetros, todas las temperaturas menos Ta y las tres X

Combinando lo aprendido en el modelo simple con la holgura y en este modelo con el aro se llega a la solución más óptima para este caso en el que se obtiene el mismo error que en el caso inicial pero con el menor número de variables consideradas.

3.1.2.2. Introducción de ruido

s1	s2	s1/s2	Tree	Linear Regression
			RMS	RMS
0,0014	0,0071	491,90	1,61E-02	1,17E-02
0,0012	0,0071	614,88	1,294E-02	9,382E-03
0,0006	0,0071	1229,76	7,142E-03	4,811E-03
0,0004	0,0071	1844,63	5,430E-03	3,322E-03
0,0003	0,0071	2459,51	4,696E-03	2,626E-03
0,0002	0,0071	3074,39	4,304E-03	2,226E-03

Tabla 8. RMSE frente a porcentaje de ruido (5)

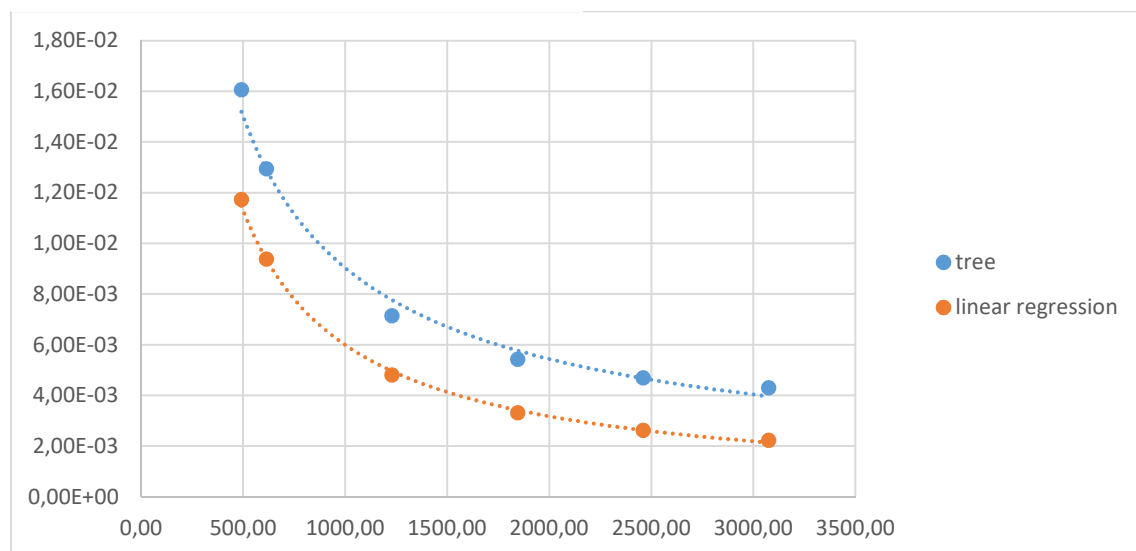


Figura 65. Representación RMSE frente a porcentaje de ruido (5)

Al igual que en todos los casos observados la tendencia seguida por la curva al aumentar el ruido introducido sigue una tendencia en la que va disminuyendo aunque cada vez a menos intensidad. En este caso se puede deducir que un error por debajo del orden de magnitud de 10^{-3} es imposible de lograr ya que en el caso en el que no se introduce ruido este no es menor a dicho orden de magnitud.

3.2. Variabilidad normal

3.2.1. Aro

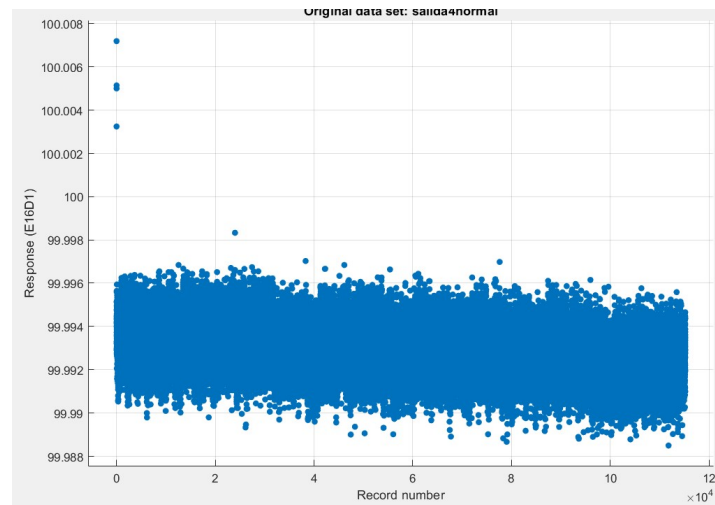


Figura 66. Diámetro frente a n (4)

Para el caso en la variabilidad normal la influencia de paso del tiempo se muestra más ligera pero sigue estando presente en los meses de verano.

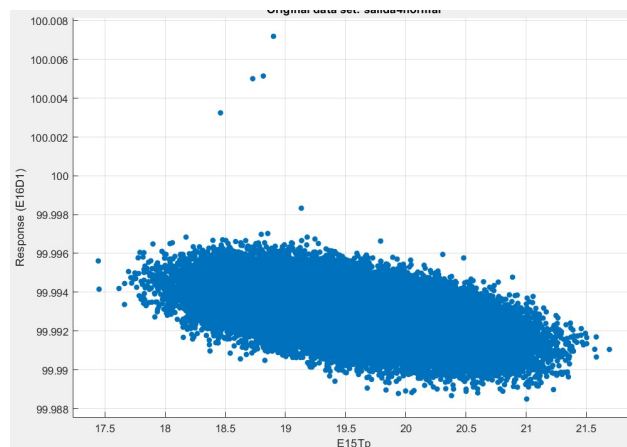


Figura 67. Diámetro frente a T (4)

Esta influencia aparece reflejada en la gráfica de la temperatura frente al diámetro donde se aprecia más claramente.

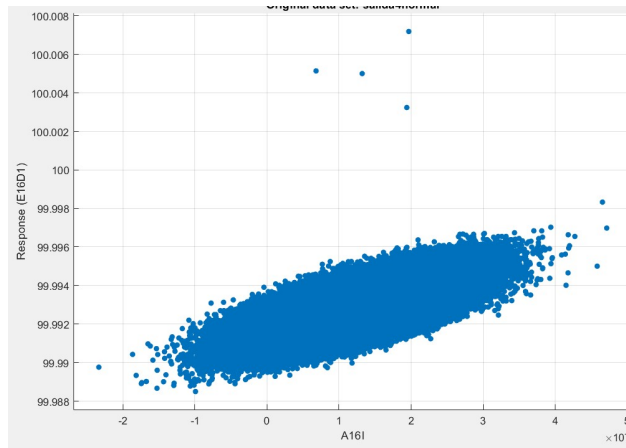


Figura 68. Diámetro frente a incertidumbre (4)

La incertidumbre también mantiene la influencia relevante frente al diámetro.

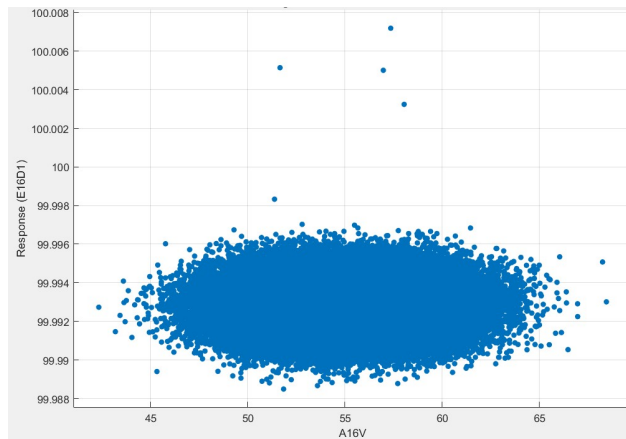


Figura 69. Diámetro frente a vibraciones (3)

El resto de variables no presentan influencia con gráficas similares a la representada de la vibración.

3.2.1.1. Cribado de factores



Figura 70. RMSE considerando todas las variables (5)

Como se ha hecho anteriormente se extrae el valor de RMSE referencial en el que se consideran todas las variables.



Figura 71. RMSE considerando la incertidumbre y todas las temperaturas menos Ta (2)

Imitando lo simulado en el modelo con variabilidad rectangular se obtiene también la simulación óptima en la que se logra el error mínimo con el mínimo número de variables.

3.2.1.2. Introducción de ruido

s1	s2	s1/s2	Tree	Linear Regression
			RMS	RMS
0,0058	0,0012	20,78	7,53E-03	5,78E-03
0,0029	0,0012	41,57	3,76E-03	2,89E-03
0,0019	0,0012	62,35	2,52E-03	1,93E-03
0,0014	0,0012	83,14	1,89E-03	1,44E-03
0,0012	0,0012	103,92	1,513E-03	1,155E-03
0,0004	0,0012	311,77	5,180E-04	3,854E-04

Tabla 9. RMSE frente a porcentaje de ruido (6)

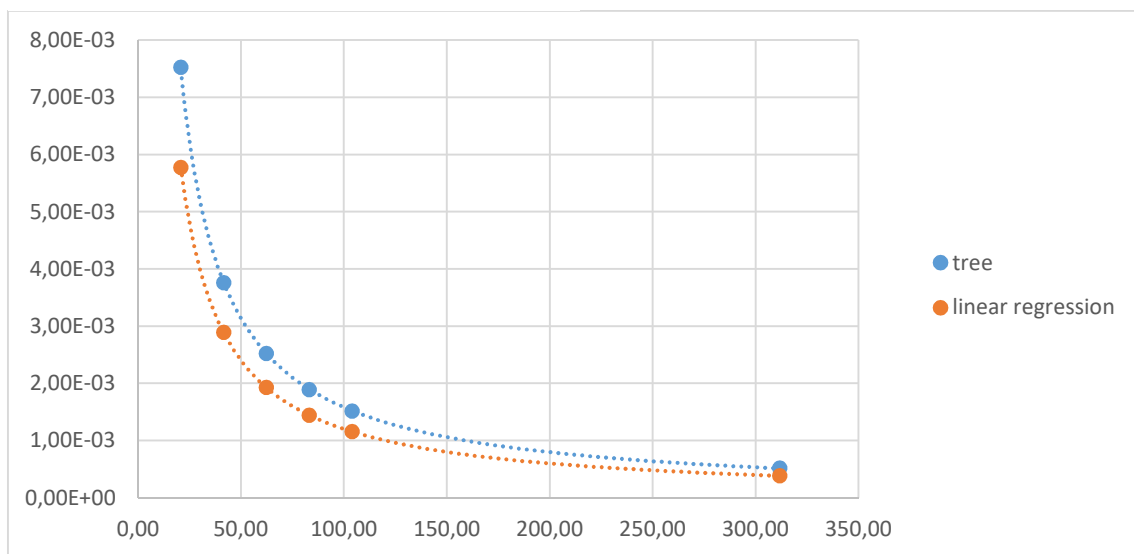


Figura 72. Representación RMSE frente a porcentaje de ruido (6)

La evolución de este modelo frente a la introducción de ruido es similar al resto en el que conforme disminuimos la intensidad de este el error cometido se va disminuyendo progresivamente cada vez menos.

3.2.2. Holgura

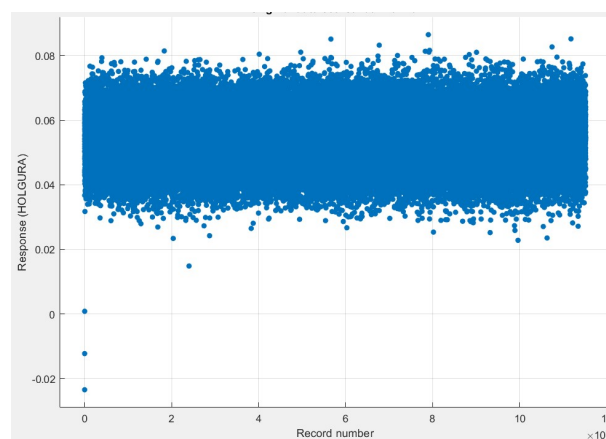


Figura 73. Holgura frente a n (4)

Al igual que con la variabilidad rectangular la holgura no se ve influenciada conforme aumenta el número de simulaciones.

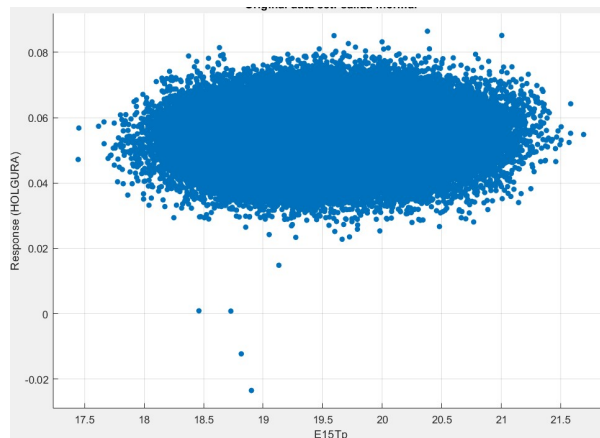


Figura 74. Holgura frente a T (4)

La nula influencia del paso del tiempo en el diámetro queda explicada ya que la temperatura tampoco presenta esa influencia.

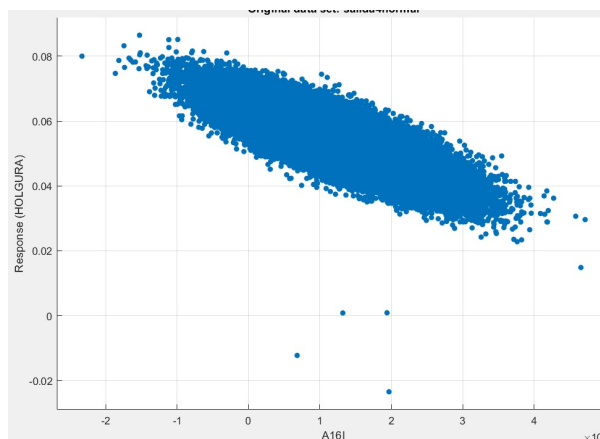


Figura 75. Holgura frente a incertidumbre (4)

La incertidumbre también sigue siendo relevante en el estudio experimental.

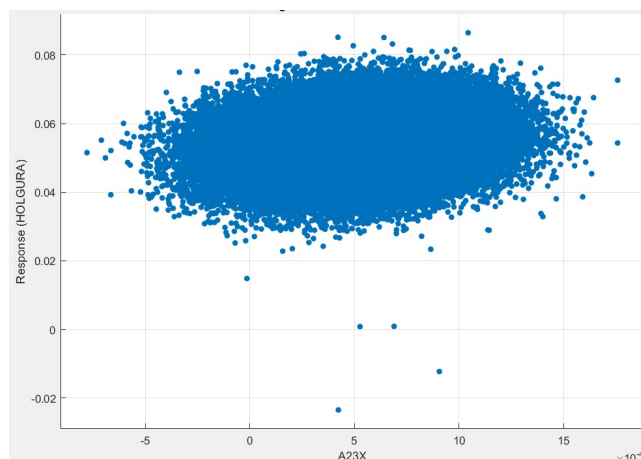


Figura 76. Holgura frente a consigna palpador (4)

Para la holgura la consigna del palpador tiene, como en casos anteriores, una ligera dependencia.

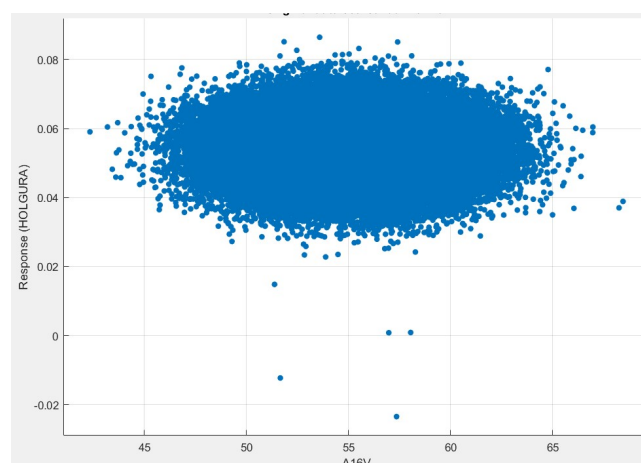


Figura 77. Holgura frente a vibración (3)

El resto de variables no presentan influencia y se podrán eliminar o no tener en cuenta en el análisis analítico posterior.

3.2.2.1. Cribado de factores



Figura 78. RMSE considerando todas las variables (6)

Se simula inicialmente la opción en la que se consideran todas las variables para comprobar el resto de simulaciones.

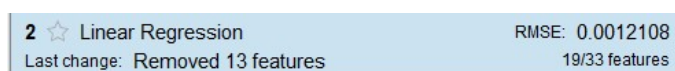


Figura 79. RMSE considerando incertidumbre, diámetros, todas las temperaturas menos Ta y las tres X (2)

Simulando la misma opción que en caso de variabilidad rectangular se obtiene la mejor opción en el que el número de variables a considerar es mínimo manteniendo el mismo error de predicción inicial.

3.2.2.2. Introducción del ruido

s1	s2	s1/s2	Tree	Linear Regression
			RMS	RMS
0,0058	0,0072	124,71	6,25E-02	4,66E-02
0,0029	0,0072	249,42	3,15E-02	2,34E-02
0,0014	0,0072	498,83	1,60E-02	1,17E-02
0,0006	0,0072	1247,08	7,22E-03	4,82E-03

Tabla 10. RMSE frente a porcentaje de ruido (7)

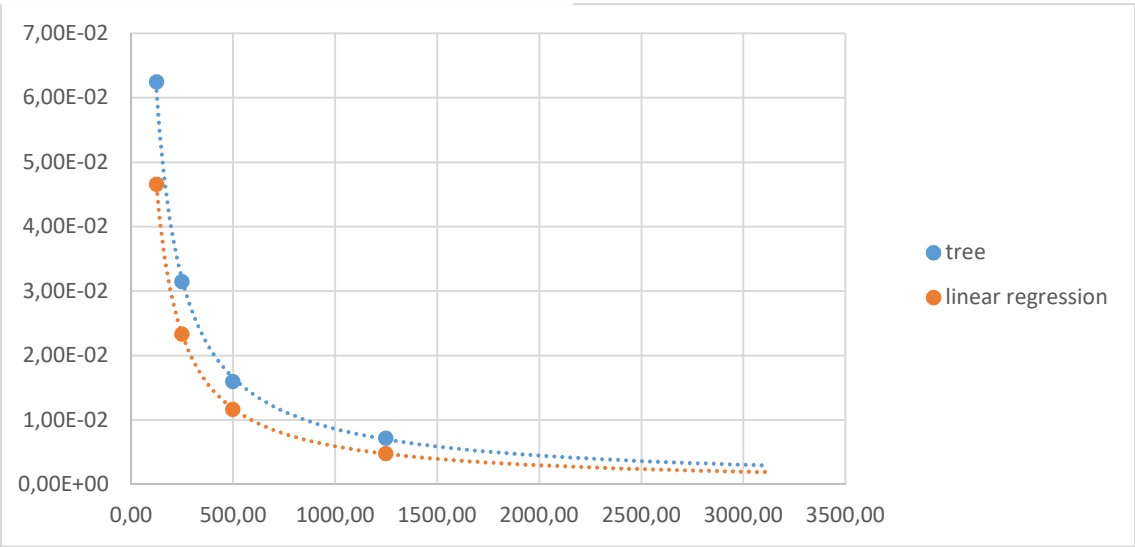


Figura 80. Representación RMSE frente a porcentaje de ruido (7)

La tendencia para este modelo es igual al de los demás estudiados. Cabe destacar que como ocurría para el estudio de la holgura en variabilidad rectangular de este modelo más complejo también resultará imposible disminuir de 10^-3 ya que para el modelo sin ruido el valor de error cometido era mayor.

4. Coste computacional

Se ha propuesto una mejora en el desarrollo de la metodología a seguir con el fin de reducir el tiempo de procesamiento si alterar el resultado de los datos obtenidos. Inicialmente de todos los procesos llevados a cabo en el trabajo el que requiere más tiempo de computación es la generación de los datos. En el Anexo1 se encuentra el código de cómo se han obtenido todos los datos inicialmente en el que se debía abrir continuamente el archivo Excel para insertar los dichos datos. Estos pasos de apertura, escritura y cierre son muy costosos computacionalmente por lo que se ha optado por almacenar dentro de una variable de tipo matriz los datos y posteriormente escribirlos en un formato TXT en el que los datos se encontrarán separados por punto y coma de esta forma se logra reducir de forma considerable el tiempo de ejecución como

se puede observar en la tabla siguiente. Este programa de generación de datos se encuentra en el Anexo2.

Generación a Excel	Generación a CSV
19 min 6 sec	12 sec

Tabla 11. Tiempos computación

El otro proceso en el que es interesante el estudio del coste computacional es el de entrenamiento de los modelos. Para este análisis se ha optado por elegir los dos métodos que inicialmente se ejecutaban de manera más rápida. Para comprobar su robustez ante la entrada masiva de datos se ha multiplicado por dos y cinco el número de variables y se ha medido el tiempo que tardan los métodos en ser entrenados. Los resultados obtenidos reflejan la superioridad computacional para este caso del modelo de *Linear Regression* frente al de *Tree*, también refleja como al aumentar el método de árbol sí que aumenta su tiempo de ejecución, sensiblemente teniendo en cuenta que se han quintuplicado los datos iniciales a procesar, mientras que el modelo de regresión consigue un aumento muy pequeño. Se comprueba como para el caso que acomete el modelo de *Linear Regression* se adapta con un error mínimo y una robustez mejor que el resto de los modelos estudiados.

	Linear Regression	Tree
Nº variables normal	41 sec	1 min 18 sec
Nº variables x 2	43 sec	2 min 1 sec
Nº variables x 5	52 sec	7 min

Tabla 12. Tiempos para diferente número de variables

5. Conclusiones y trabajo futuro

Como se ha explicado el proceso de producción de los rodamientos no se podrá analizar mediante métodos estadísticos por lo que una solución como la de la implementación de métodos de Machine Learning resulta atractiva para llegar a una predicción idónea.

Se han demostrado las dependencias existentes que aparecen entre las variables y las dimensiones finales y cómo la supresión de alguna de estas no suponen un aumento en el error por lo que será irrelevante su consideración. También se han propuesto los modelos de predicción más exactos para cada modelo así como su comportamiento frente a diferentes ruidos así como una mejora en el programa de generación de datos para disminuir su tiempo de trabajo. Por último se ha demostrado la robustez de los modelos seleccionados midiendo sus tiempos de compilación y cómo dichos modelos reaccionan al aumento de variables, siendo crucial la reducción del número de estas.

Como ampliación al TFG se debería exportar el modelo entrenado desarrollado con la aplicación de Matlab al lenguaje de Python. Esto permitirá el desarrollo y la interacción con diversas APIs las cuales nutrirán al modelo de nuevos datos conforme el proceso avance así como la creación de una base de datos donde almacenar todos los datos registrados y el desarrollo de un modelo ETL (Extract, Transform, Load) el cual extraiga los datos de los repositorios de generación, los transforme adecuando los datos para el posterior uso por parte del modelo de regresión y finalmente los cargue en las bases de datos de donde se nutrirá dicho modelo. También esto

permitirá interaccionar con las maquinas pudiendo dar órdenes de cambio en las variables de producción en tiempo real permitiendo el deseado cero errores.

Otro trabajo posterior es el estudio de la posibilidad de incluir un análisis PCA en los modelos matemáticos. Este análisis ayuda en la reducción del número de variables y además previene el “overfitting” o sobreajuste el cual impide que el método generalice provocando mayor error de predicción. El análisis PCA lo consigue mediante la eliminación de las variables que no son influyentes como se ha realizado de forma manual y mediante la creación de nuevas variables combinación de las iniciales. De esta forma mejora la eficiencia del modelo siendo capaz de reducir el número de variables a unas pocas con gran influencia sobre el resultado final. Esta eficiencia provoca que las variables generadas por el análisis sean variables abstractas o combinación de ellas lo que hace perder el sentido del significado de las variables.

El estudio de los procesos de fabricación mediante Machine Learning así como la recopilación de datos consigue, mediante una utilización inteligente de la información, conocer mejor el proceso productivo pudiendo reducir errores y reduciendo los tiempos.

Anexo 1. Código generación de datos inicial

```
close all;
clc;
warning off all;
tic;

filename = 'salida3.xlsx';
n=1e5; % máximo 1048576 antes 65636
piezas_hora=40;
% dias=20*12; %máximo!
dias=1*1;
n=dias*piezas_hora*24;

% xx1=2*(randn(n,1)); %normal
% xx1=(rand(n,1))-0.5; %rectangular

% figure
% hist(xx1,400)

Tref = NaN(n,1);
Tref = gen_temperatura(dias,piezas_hora);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ARO
xx1=(rand(n,1))-0.5; %Ta %rectangular [-0.5,0.5]
xx2=(rand(n,1))-0.5; %Tt
xx3=(rand(n,1))-0.5; %Tx
xx4=(rand(n,1))-0.5; %Tp antes
xx5=(rand(n,1))-0.5; %Tp despues
xx6=(rand(n,1))-0.5; %diametro previo
xx7=((rand(n,1))-0.5)/10;
xx8=((rand(n,1))-0.5)/10;

figure
hist(xx7,400)
ssl=std(xx7)
```

```

A1_6_Ta=20+(Tref-20)/5+xx1;
col='C';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_Ta,str1);

```

```

A1_6_V=50+rand(n,1)*10;
col='D';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_V,str1);

```

```

A1_6_R=60+rand(n,1)*20;
col='E';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_R,str1);

```

```

%A1_6_X=;%%%%%%%%%%%%%%|
A1_6_EX=rand(n,1)/1000;
col='F';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_EX,str1);

```

```

A1_6_I=rand(n,1)/400;
col='G';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_I,str1);

```

```

A1_6_Tt=20+(Tref-20)/10+xx2;
col='H';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_Tt,str1);

```

```

A1_6_Tx=20+(Tref-20)/15+xx3;
col='I';
str1=sprintf('%s6:%s%d',col,col,n+5);

```

```

xlswrite(filename,A1_6_P,str1);

A1_5_Tp=20+(Tref-20)/12+xx4;
col='K';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_5_Tp,str1);

A1_6_Tp=20+(Tref-20)/5+10+xx5;
col='L';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_Tp,str1);

A1_5_Dl=100+xx6/500;
col='M';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_5_Dl,str1);

A1_6_Dl=A1_5_Dl.*(1+11.2e-6*(20-A1_6_Tp))+A1_6_I ; %mejorar con temperaturas
col='N';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,A1_6_Dl,str1);
figure
hist(A1_6_Dl,400)
ss3=std(A1_6_Dl)

HOLGURA=(A3_3_d1-100)*4+(A2_3_d1-100)*4-(A1_6_Dl-100)*8+(rand(n,1)-0.5)*4/1000+0.060;
col='AM';
str1=sprintf('%s6:%s%d',col,col,n+5);
xlswrite(filename,HOLGURA,str1);
toc;
figure
hist(HOLGURA,400)
ss4=std(HOLGURA)

%-----
function [T] = gen_temperatura(D, P)
    pp=1;
    filename = 'temperatura.xlsx';
    col='A';
    str=sprintf('%s1:%s%d',col,col,D);
    Tmax = xlsread(filename,'Hoja2',str);
    col='B';
    str=sprintf('%s1:%s%d',col,col,D);
    Tmin = xlsread(filename,'Hoja2',str);
    col='C';
    str=sprintf('%s1:%s%d',col,col,24);
    Thora = xlsread(filename,'Hoja2',str);
    for dia=1:D
        for hora=1:24
            for pieza=1:P
                T(pp,1)=Thora(hora)+(Tmin(dia)-13)+(Tmax(dia)-Tmin(dia))/(26-13);
                pp=pp+1;
            end
        end
    end
end

```

Anexo 2. Código de generación mejorado

```
|close all;
clc;
warning off all;
tic;

% factor = 150; filename = 'salidaAño4normal_150.txt'; %s/r=2.7 (600)
% factor = 125; filename = 'salidaAño4normal_125.txt'; %s/r=2.4 (500)
factor = 100; filename = 'salidaAño4normal_100.txt'; %s/r=2.0 (400)
% factor = 75; filename = 'salidaAño4normal_75.txt'; %s/r=1.6 (300)
% factor = 50; filename = 'salidaAño4normal_50.txt'; %s/r=1.3 (200)
% factor = 10; filename = 'salidaAño4normal_10.txt'; %s/r=1.0 ( 40)

factor_ruido = 400*factor/100;

%n=1e5; % máximo en Excel 1048576 antes 65636
piezas_hora=40;
% dias=20*12; %máximo!
%dias=1*1; % 1600
%dias=20*1;% 19200
dias=20*12;% 384000
n=dias*piezas_hora*24;

% xx1=0.5/3^0.5*(randn(n,1)); %normal s=0.5/3^0.5 o 0.2886
% xx2=(rand(n,1))-0.5; %rectangular [-0.5,0.5]
% rand(n,1) equivale a (0.5/3^0.5*(randn(n,1))+0.5)
% xx1=50+(0.5/3^0.5*(randn(n,1))+0.5)*10; %normal
% xx2=50+rand(n,1)*10; %rectangular
% figure
% hist(xx1,400)
% figure
% hist(xx2,400)

Tref = NaN(n,1);
Tref = gen temperatura(dias,piezas_hora);
```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% ARO
xx1=0.5/3^0.5*(randn(n,1)); %Ta
xx2=0.5/3^0.5*(randn(n,1)); %Tt
xx3=0.5/3^0.5*(randn(n,1)); %Tx
xx4=0.5/3^0.5*(randn(n,1)); %Tp antes
xx5=0.5/3^0.5*(randn(n,1)); %Tp despues
xx6=0.5/3^0.5*(randn(n,1)); %diametro previo
senal=0*(randn(n,1)); %para info señal/ruido

A1_6_Ta=20+(Tref-20)/5+xx1;
A1_6_V=50+(0.5/3^0.5*(randn(n,1))+0.5)*10;
A1_6_R=60+(0.5/3^0.5*(randn(n,1))+0.5)*20;
%A1_6_X=;%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
A1_6_EX=(0.5/3^0.5*(randn(n,1))+0.5)/1000;
A1_6_I=(0.5/3^0.5*(randn(n,1))+0.5)/400;
S1_6_Tt=20+(Tref-20)/10+xx2;
S1_6_Tx=20+(Tref-20)/15+xx3;
S1_6_P=50+(0.5/3^0.5*(randn(n,1))+0.5)*5;
E1_5_Tp=20+(Tref-20)/12+xx4;
E1_6_Tp=20+(Tref-20)/5+10+xx5;

E1_5_D1=100+xx6/500;
ruido_D1=(0.5/3^0.5*(randn(n,1))+0.5)/factor_ruido;

%E1_6_D1=A1_5_D1.*(1+11.2e-6*(20-A1_6_Tp))+A1_6_I; %3 mejorar con temperaturas
E1_6_D1=(mean(E1_5_D1)*(1+11.2e-6*(20-(E1_5_Tp*0.2+E1_6_Tp*0.8)))).*(1+11.2e-6*((S1_6_Tt*0.2+S1_6_Tx*0.8)-20))
senal(1)=std(E1_6_D1);
senal(2)=std(ruido_D1);
senal(3)=senal(1)/senal(2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% EMPAREJADO

ruido_HOLGURA=(0.5/3^0.5*(randn(n,1))+0.5)/factor_ruido*7.7;

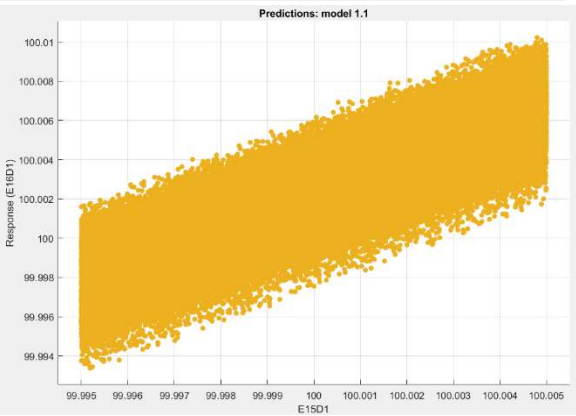
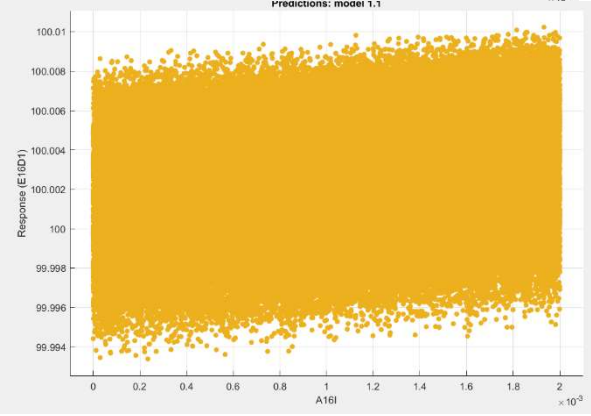
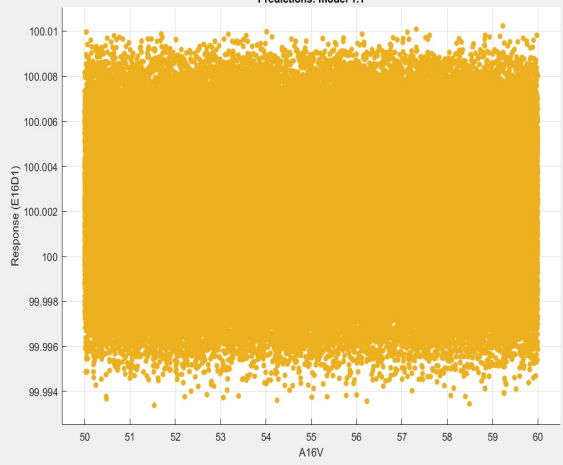
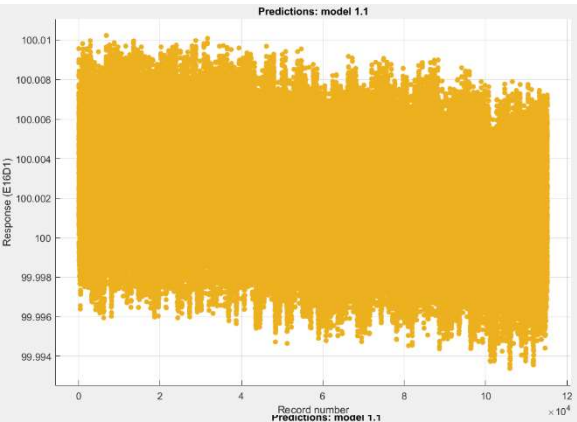
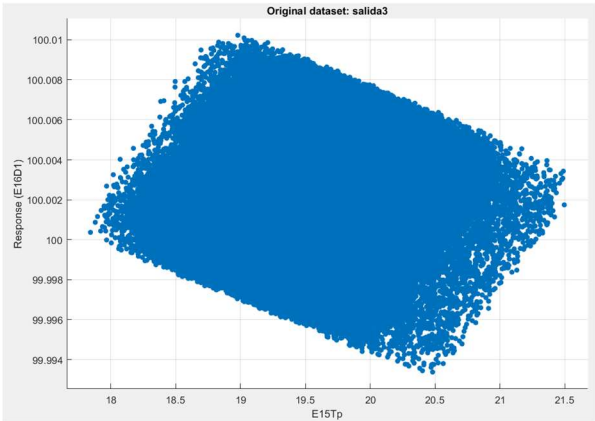
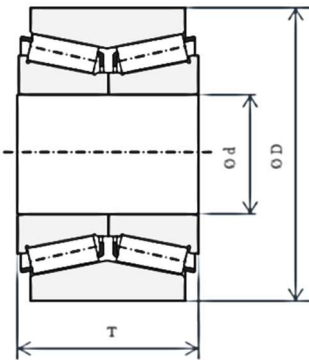
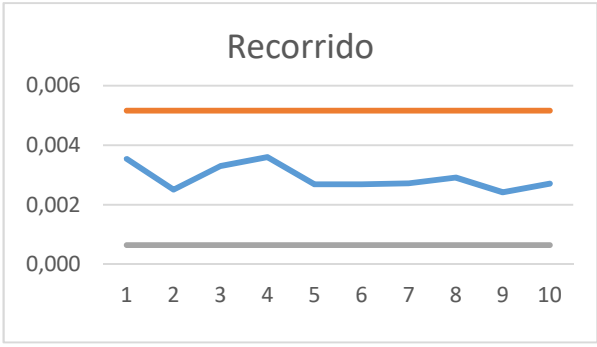
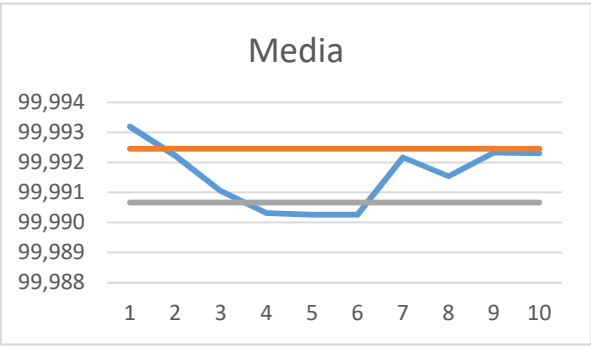
HOLGURA=(E3_3_d1-100)*4+(E2_3_d1-100)*4-(E1_6_D1-100)*8+((0.5/3^0.5*(randn(n,1))+0.5)-0.5)*4/1000+0.060+ruido_HOLGURA;
senal(13)=std(HOLGURA);
senal(14)=std(ruido_HOLGURA);
senal(15)=senal(13)/senal(14);

T=table(senal,A1_6_Ta,A1_6_V,A1_6_R,A1_6_EX,A1_6_I,S1_6_Tt,S1_6_Tx,S1_6_P,E1_5_Tp,E1_6_Tp,E1_5_D1, E1_6_D1,A2_3_Ta,
A2_3_V,A2_3_R,A2_3_EX,A2_3_D,S2_3_Tt,S2_3_Tx,S2_3_P,E2_2_Tp,E2_3_Tp,E2_2_d1,E2_3_d1,A3_3_Ta,A3_3_V,A3_3_R,A3_3_EX,
A3_3_D,S3_3_Tt,S3_3_Tx,S3_3_P,E3_2_Tp,E3_3_Tp,E3_2_d1,E3_3_d1,HOLGURA);
writetable(T,filename);

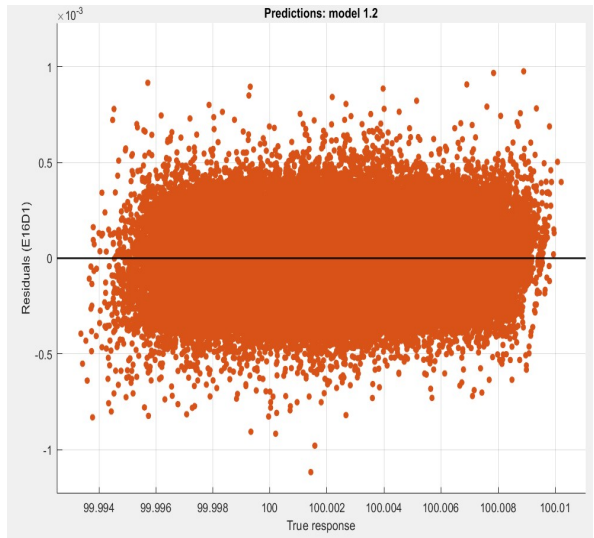
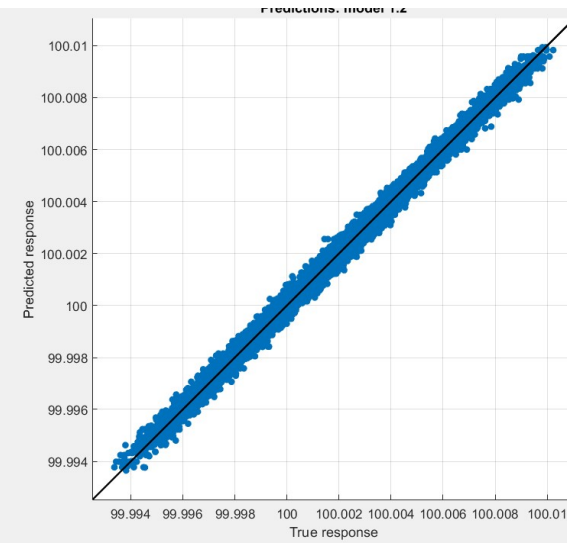
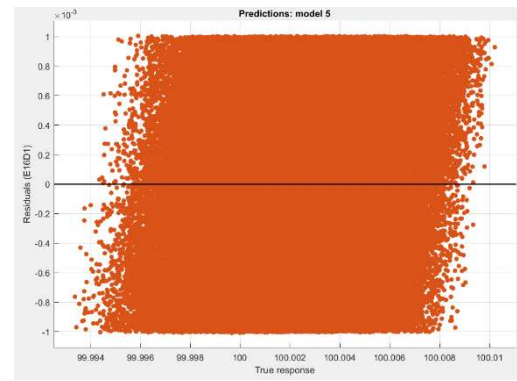
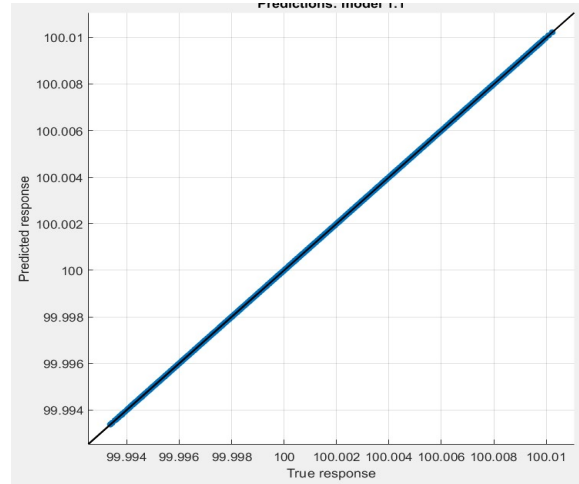
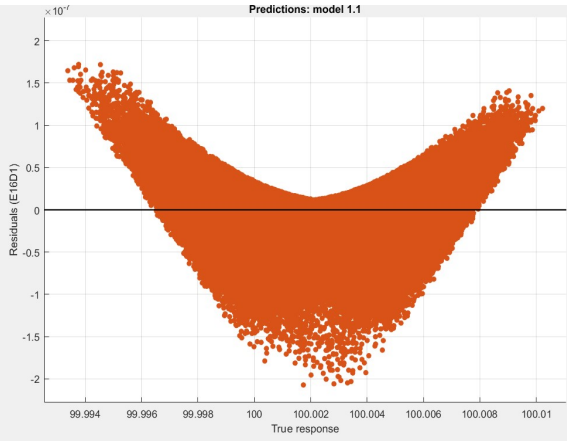
%-----
function [T] = gen_temperatura(D, P)
    pp=1;
    filename = 'temperatura.xlsx';
    col='A';
    str=sprintf('%s1:%s%d',col,col,D);
    Tmax = xlsread(filename,'Hoja2',str);
    col='B';
    str=sprintf('%s1:%s%d',col,col,D);
    Tmin = xlsread(filename,'Hoja2',str);
    col='C';
    str=sprintf('%s1:%s%d',col,col,24);
    Thora = xlsread(filename,'Hoja2',str);
    for dia=1:D
        for hora=1:24
            for pieza=1:P
                T(pp,1)=Thora(hora)+(Tmin(dia)-13)+(Tmax(dia)-Tmin(dia))/(26-13);
                pp=pp+1;
            end
        end
    end
end

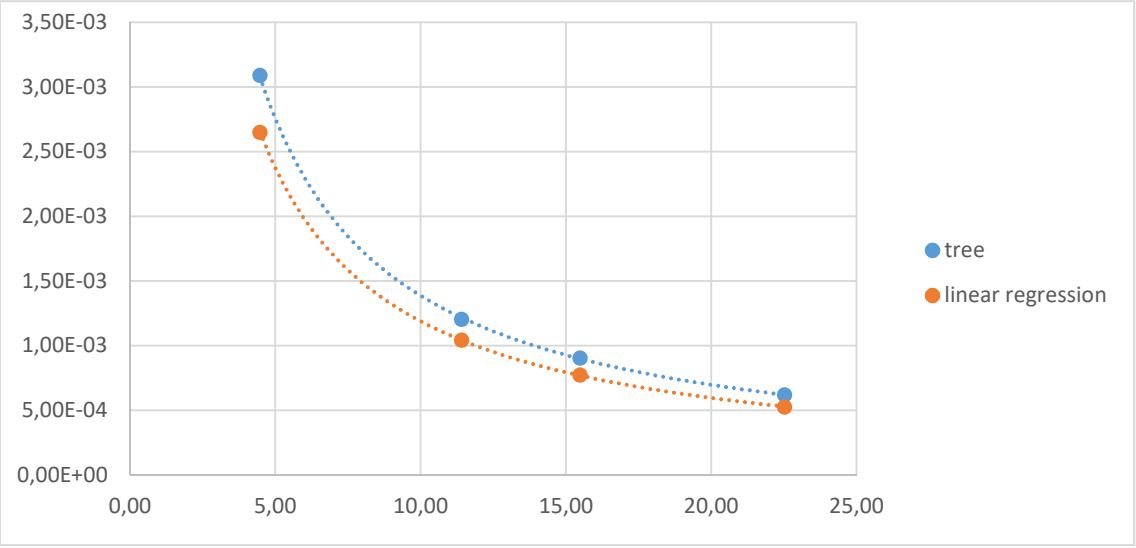
```


Anexo 3. Análisis con Regression Learner de Matlab

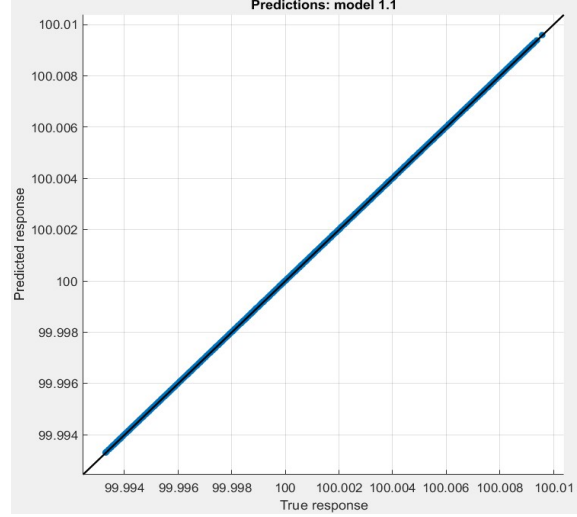


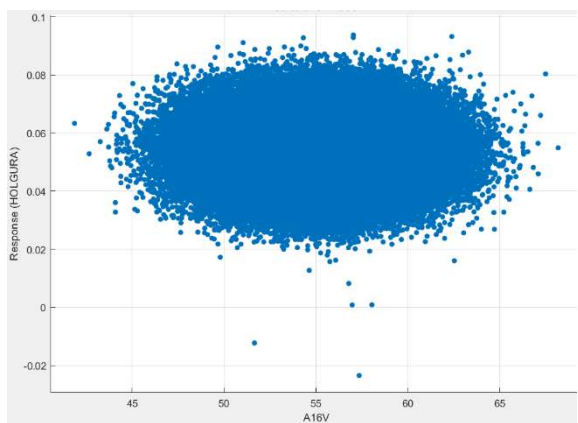
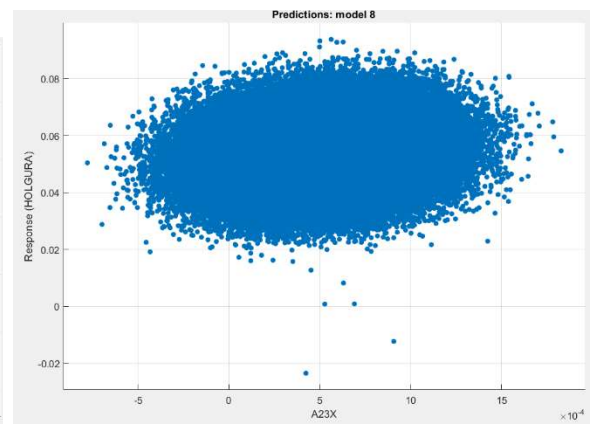
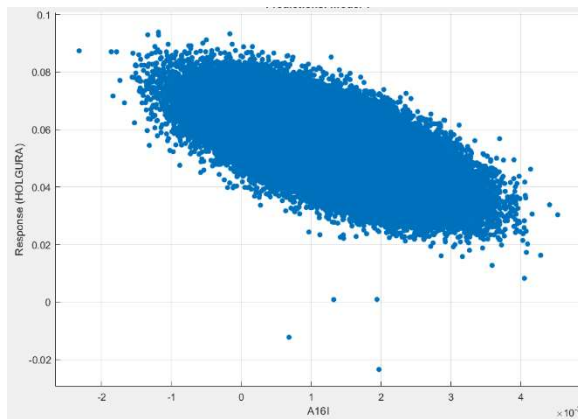
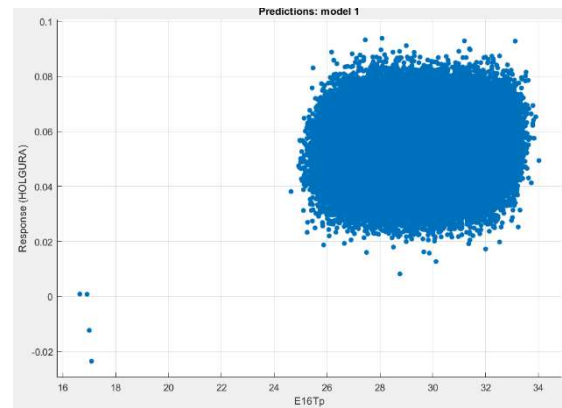
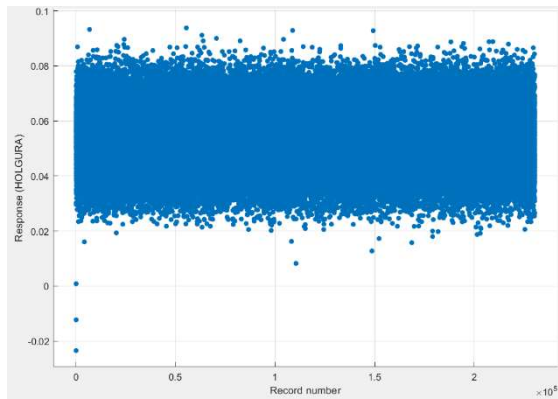
1.1 ☆ Linear Regression	RMSE: 0.00
Last change: Linear	11/11 features
1.2 ☆ Tree	RMSE: 0.00
Last change: Complex Tree	11/11 features
1.3 ☆ Tree	RMSE: 0.00
Last change: Medium Tree	11/11 features
1.4 ☆ Tree	RMSE: 0.00
Last change: Simple Tree	11/11 features





1.1	☆ Linear Re...	RMSE: 0.00
	Last change: Linear	11/11 features
1.2	☆ Tree	RMSE: 0.00
	Last change: Compl...	11/11 features
1.3	☆ Tree	RMSE: 0.00
	Last change: Mediu...	11/11 features
1.4	☆ Tree	RMSE: 0.00
	Last change: Simple...	11/11 features



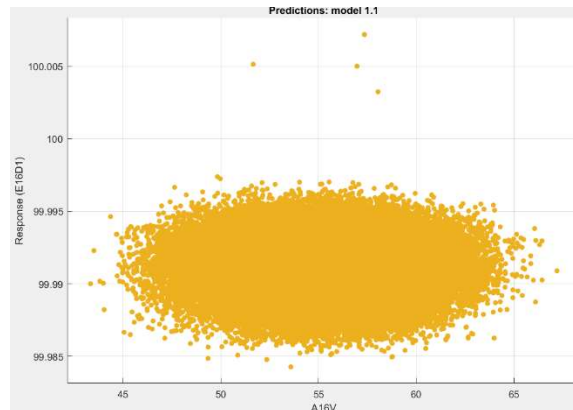
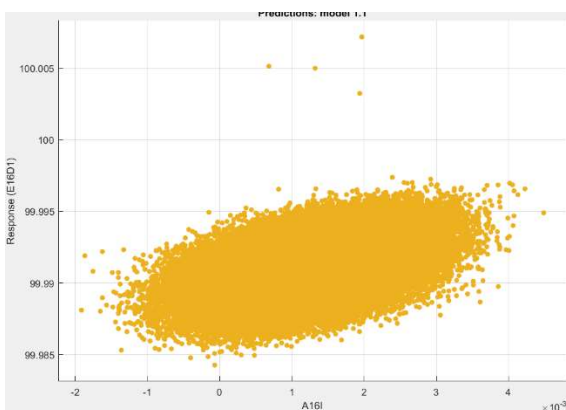
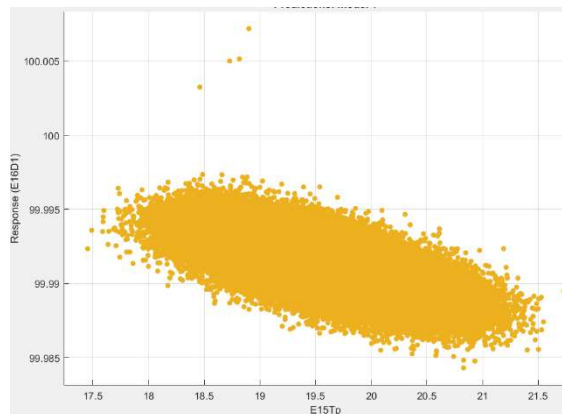
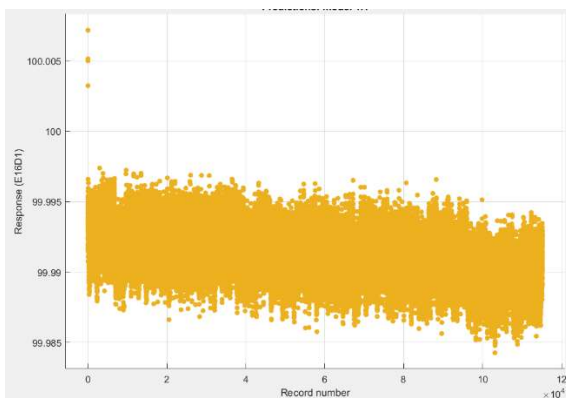
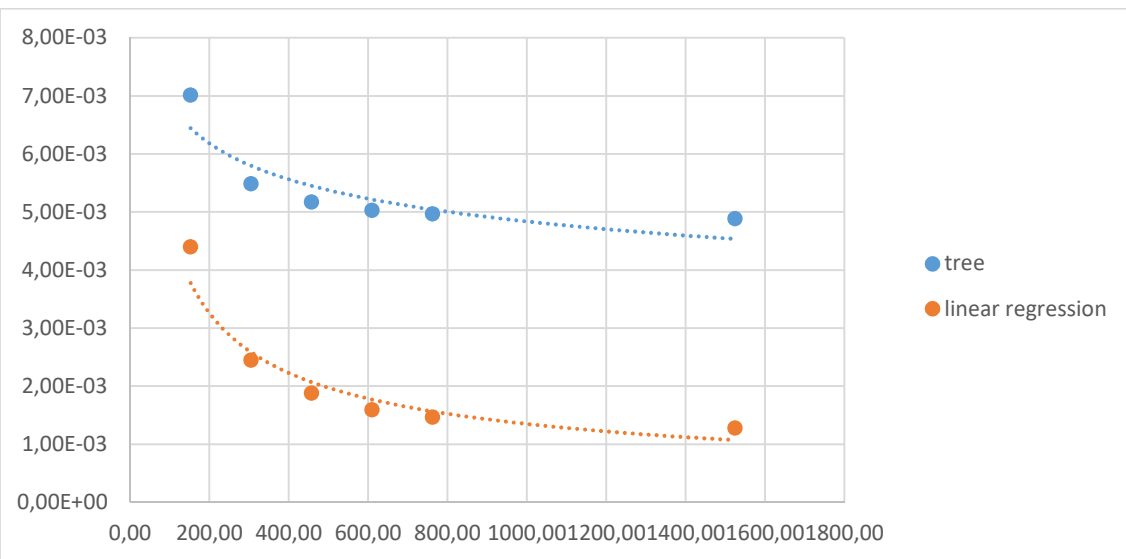


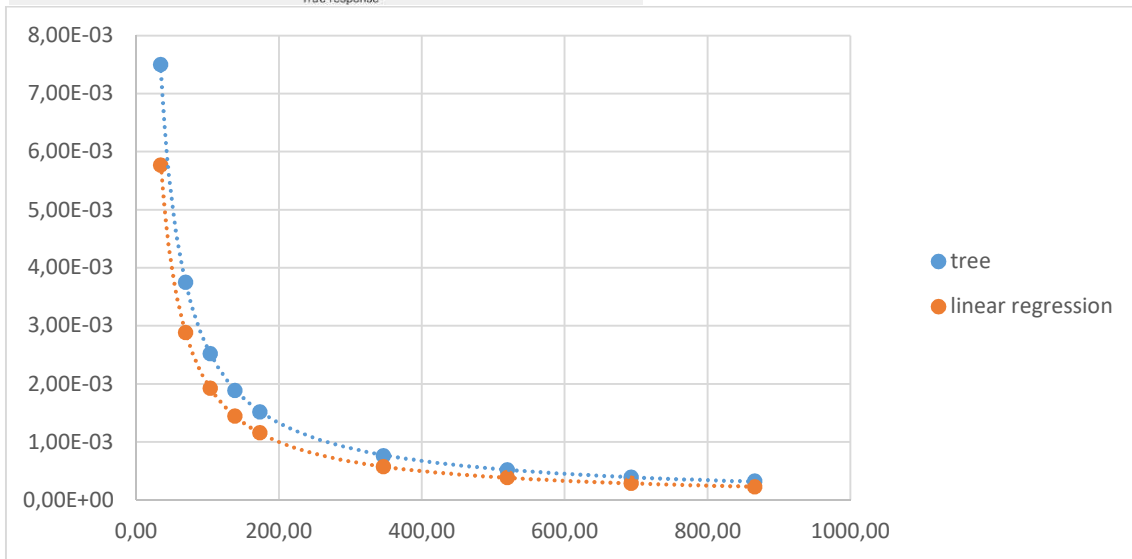
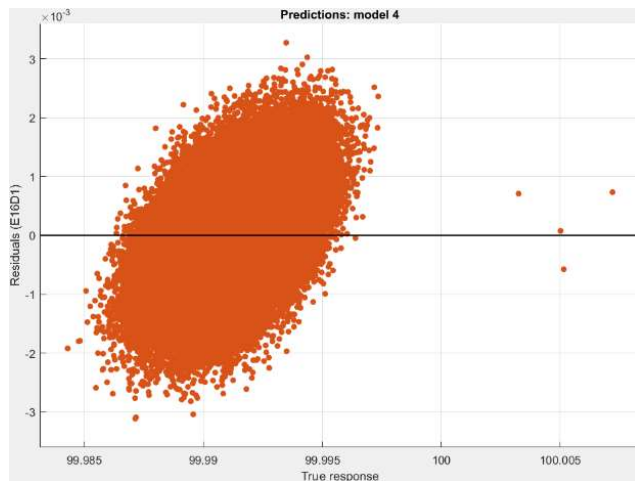
1 ☆ Linear Regression RMSE: **0.0011824**
Last change: Linear 33/33 features

2 ☆ Linear Regression RMSE: 0.0037505
Last change: Disabled PCA 4/33 features

7 ☆ Linear Regression RMSE: 0.0020148
Last change: Changed 5 features 7/33 features

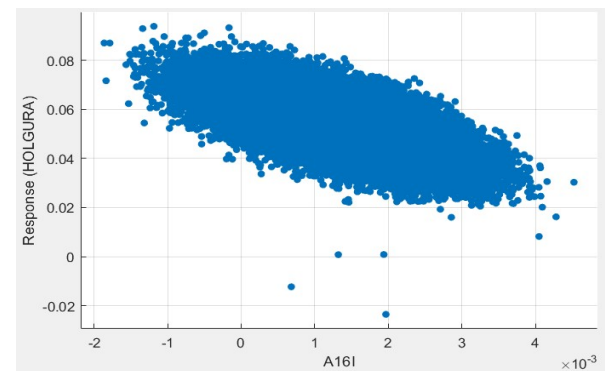
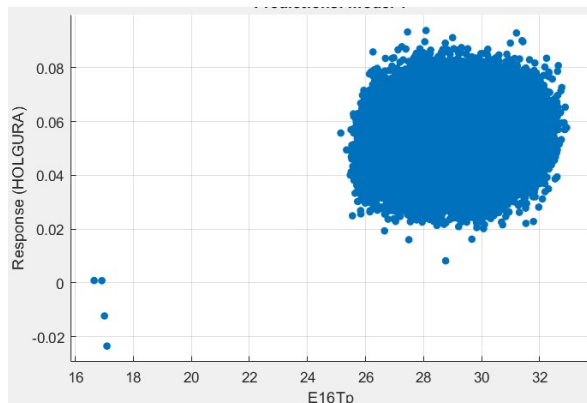
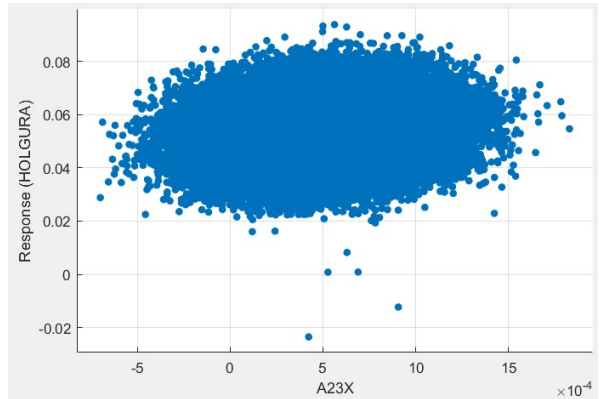
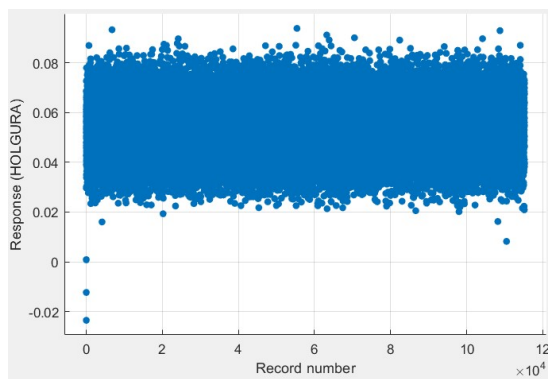
8 ☆ Linear Regression	RMSE: 0.0037435
Last change:	6/33 features
6 ☆ Linear Regression	RMSE: 0.0011827
Last change: Added 2 features	10/33 features

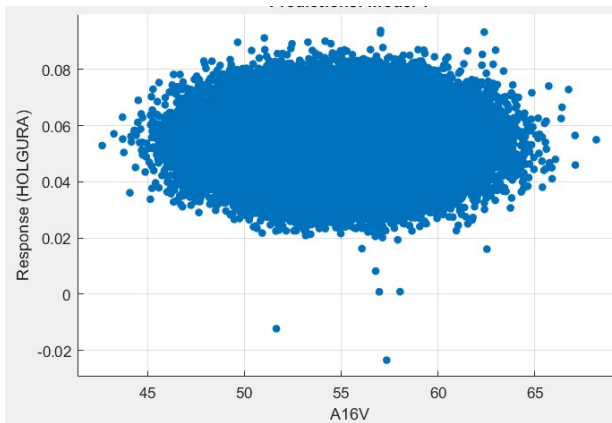




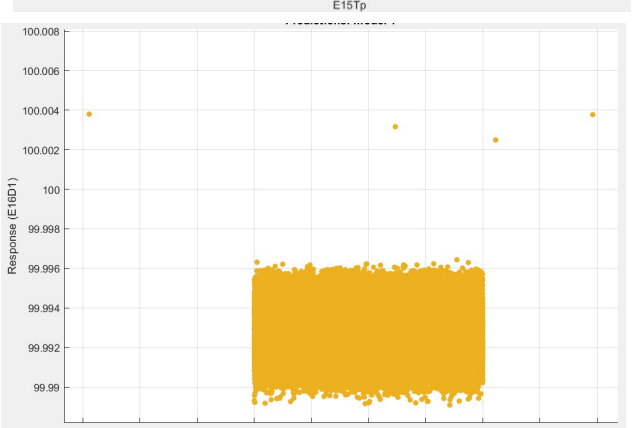
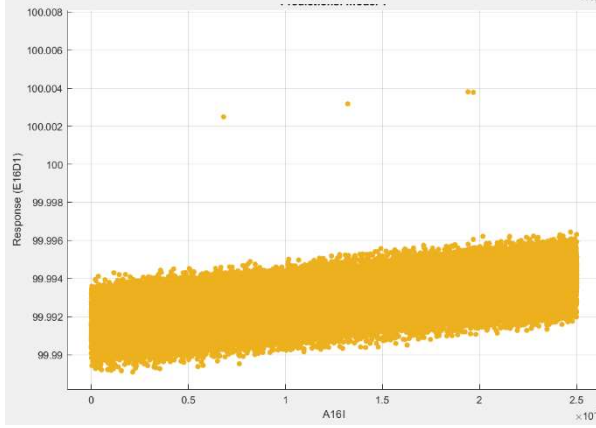
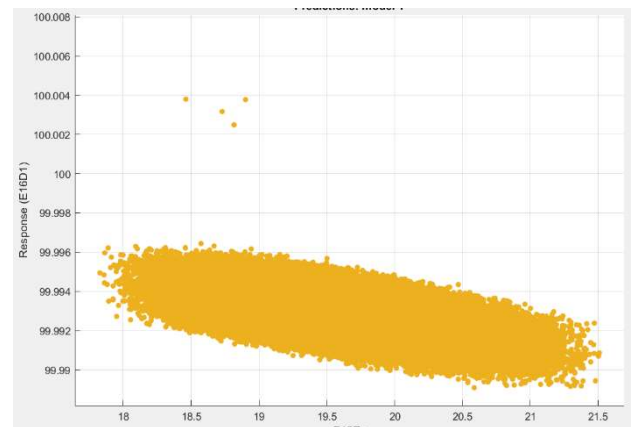
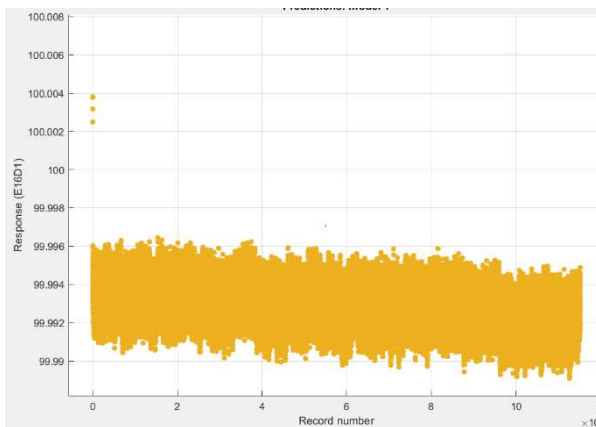
1.1 ☆ Linear Regression	RMSE: 0.0057725
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0057737
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.0057725
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0057732
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0075027
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.0066863
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0061336
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.0057742
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0057781
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0057904
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0060315
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0058232
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0057776
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2387
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0058606
Last change: Bagged Trees	11/11 features

1.1 ☆ Linear Regression	RMSE: 0.001159
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0011595
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.001159
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0011591
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0015187
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.0013504
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0012411
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.0011607
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0011634
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0011648
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0015022
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0011751
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0011637
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2387
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0011872
Last change: Bagged Trees	11/11 features

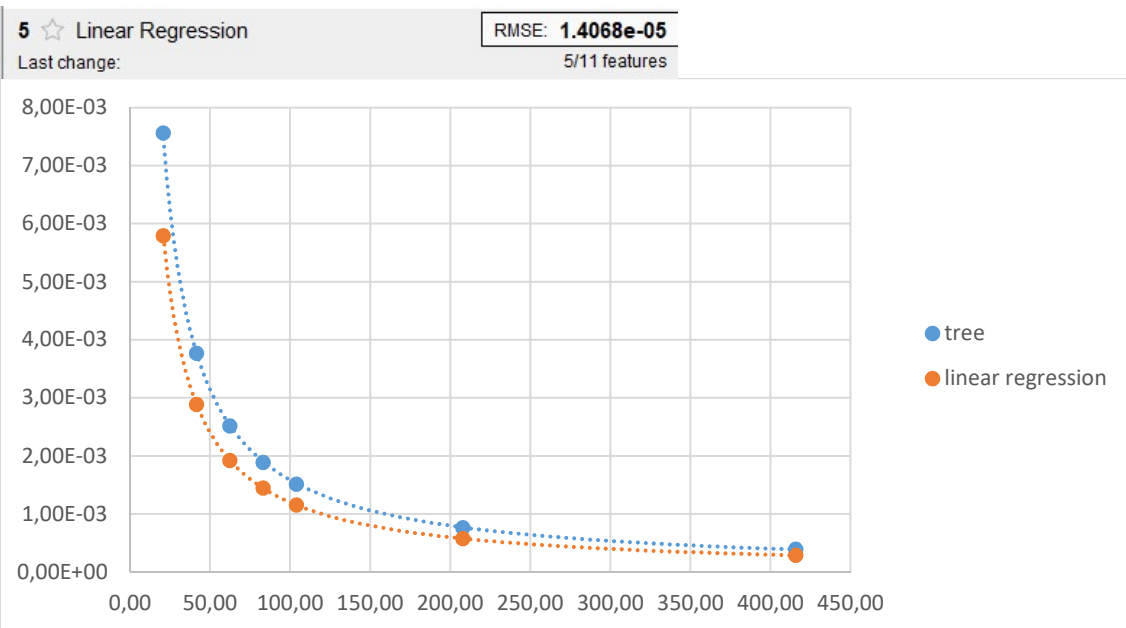




1	☆ Linear Regression	RMSE: 0.0012052
Last change: Linear		33/33 features
2	☆ Linear Regression	RMSE: 0.0012069
Last change: Removed 22 features		10/33 features
3	☆ Linear Regression	RMSE: 0.0020317
Last change: Removed 2 features		7/33 features
6	☆ Linear Regression	RMSE: 7.2735e-05
Last change:		4/11 features

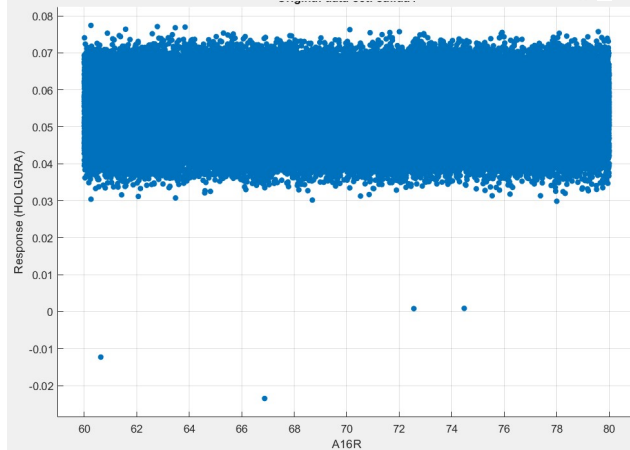
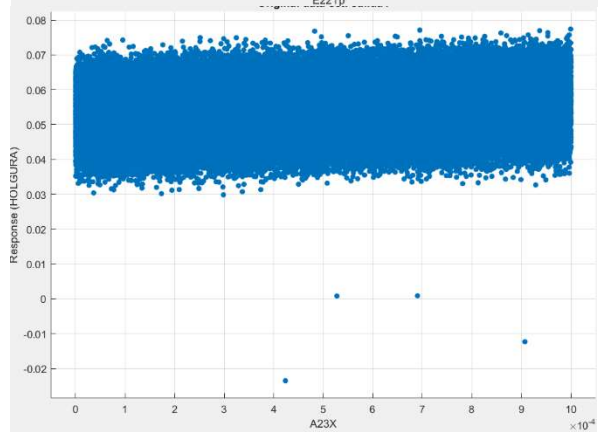
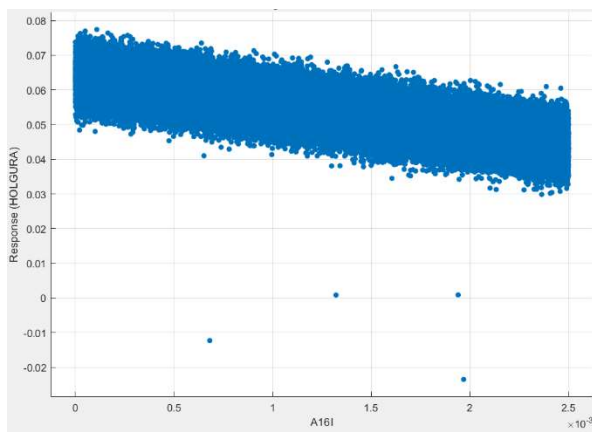
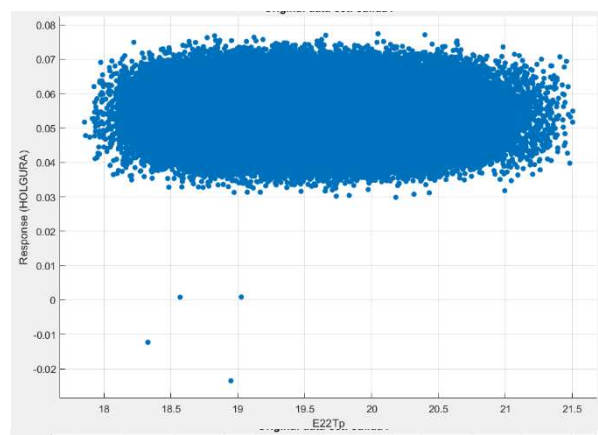
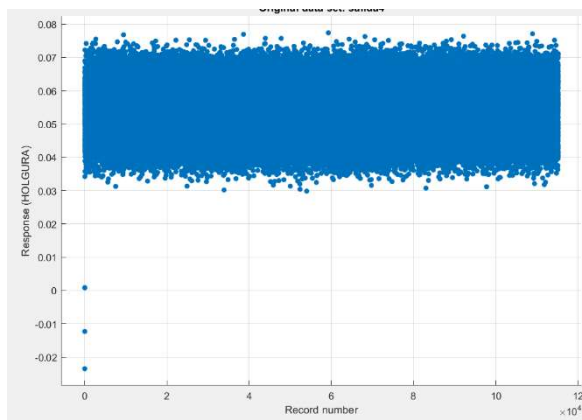


1	☆ Linear Regression	RMSE: 1.4087e-05
Last change: Linear		11/11 features
2	☆ Linear Regression	RMSE: 0.00029078
Last change: Removed 8 features		2/11 features
4	☆ Linear Regression	RMSE: 1.4086e-05
Last change: Added 4 features		6/11 features

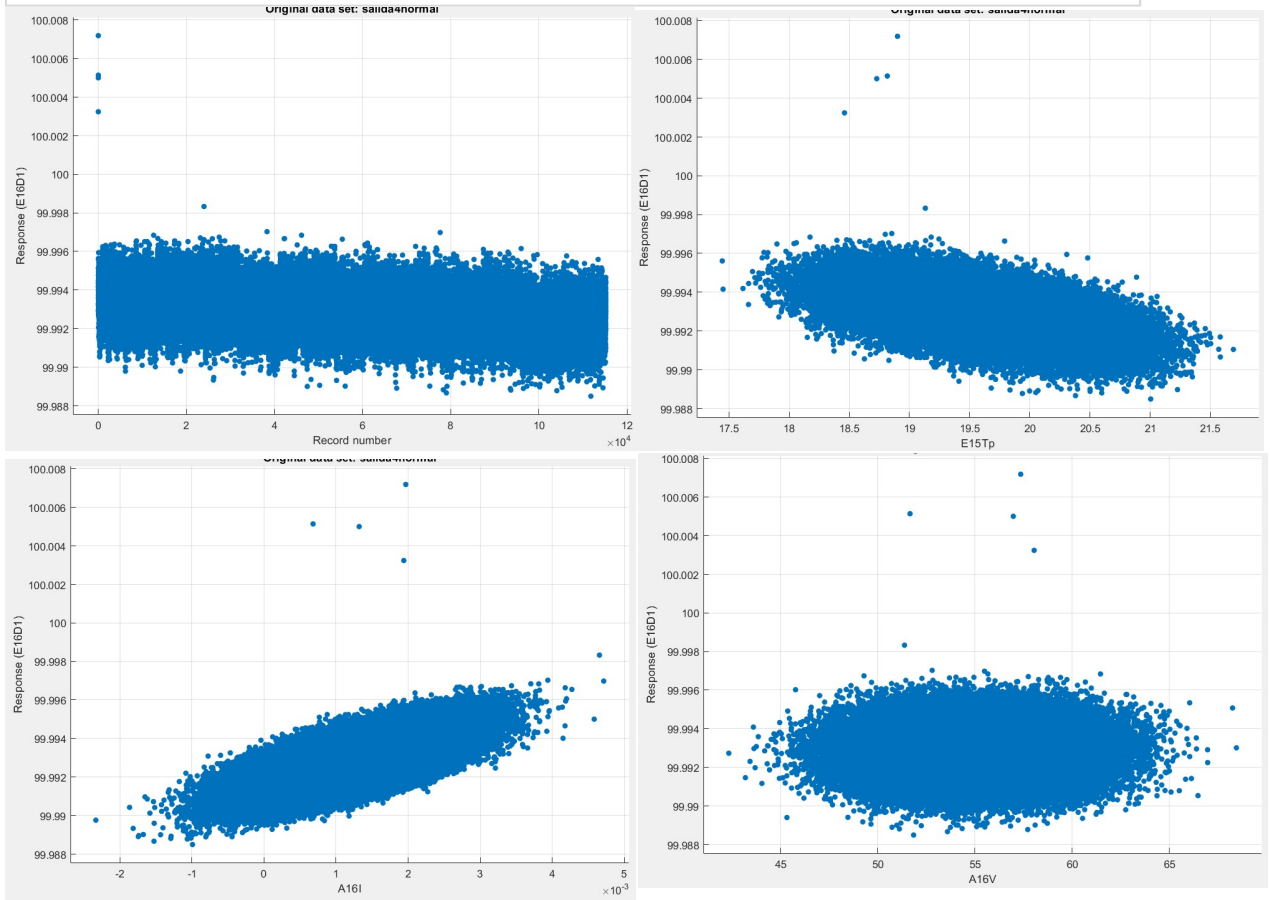
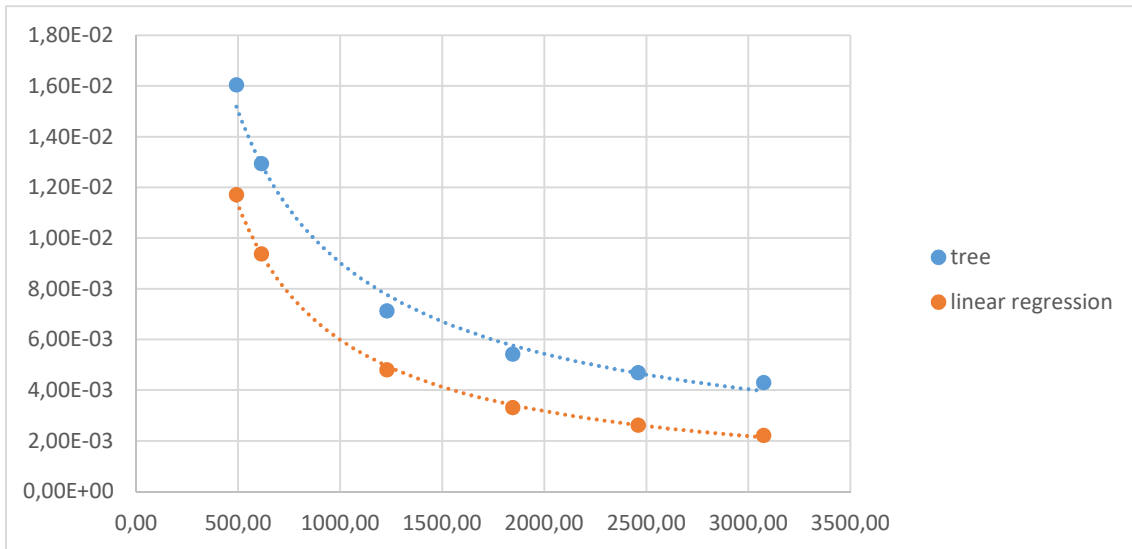


1.1 ☆ Linear Regression	RMSE: 0.0011541
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0011545
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.0011542
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0011543
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0015156
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.0013473
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0012316
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.001157
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0011603
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0011632
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0013014
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0011733
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0011583
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2388
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0011779
Last change: Bagged Trees	11/11 features

1.1 ☆ Linear Regression	RMSE: 0.0057892
Last change: Linear	11/11 features
1.2 ☆ Linear Regression	RMSE: 0.0057912
Last change: Interactions Linear	11/11 features
1.3 ☆ Linear Regression	RMSE: 0.0057892
Last change: Robust Linear	11/11 features
1.4 ☆ Stepwise Linear Regression	RMSE: 0.0057894
Last change: Stepwise Linear	11/11 features
1.5 ☆ Tree	RMSE: 0.0075624
Last change: Fine Tree	11/11 features
1.6 ☆ Tree	RMSE: 0.006729
Last change: Medium Tree	11/11 features
1.7 ☆ Tree	RMSE: 0.0061531
Last change: Coarse Tree	11/11 features
1.8 ☆ SVM	RMSE: 0.0057914
Last change: Linear SVM	11/11 features
1.9 ☆ SVM	RMSE: 0.0058018
Last change: Quadratic SVM	11/11 features
1.10 ☆ SVM	RMSE: 0.0058197
Last change: Cubic SVM	11/11 features
1.11 ☆ SVM	RMSE: 0.0061004
Last change: Fine Gaussian SVM	11/11 features
1.12 ☆ SVM	RMSE: 0.0058709
Last change: Medium Gaussian SVM	11/11 features
1.13 ☆ SVM	RMSE: 0.0057982
Last change: Coarse Gaussian SVM	11/11 features
1.14 ☆ Ensemble	RMSE: 4.2388
Last change: Boosted Trees	11/11 features
1.15 ☆ Ensemble	RMSE: 0.0058692
Last change: Bagged Trees	11/11 features

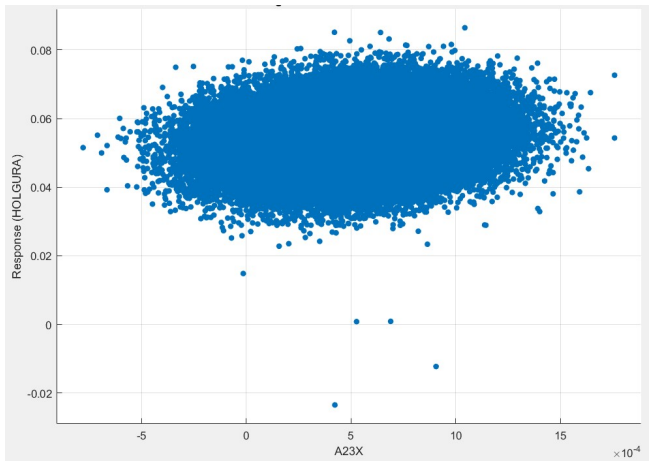
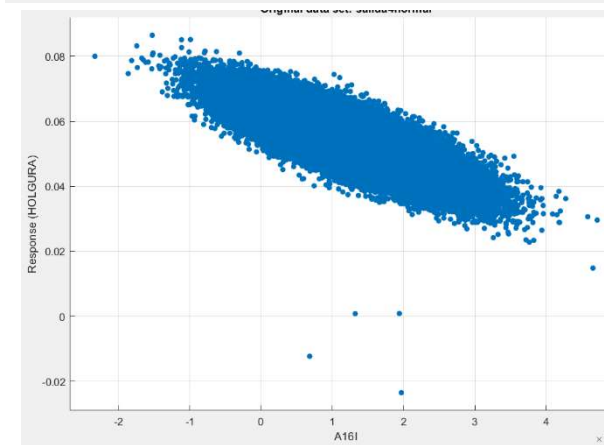
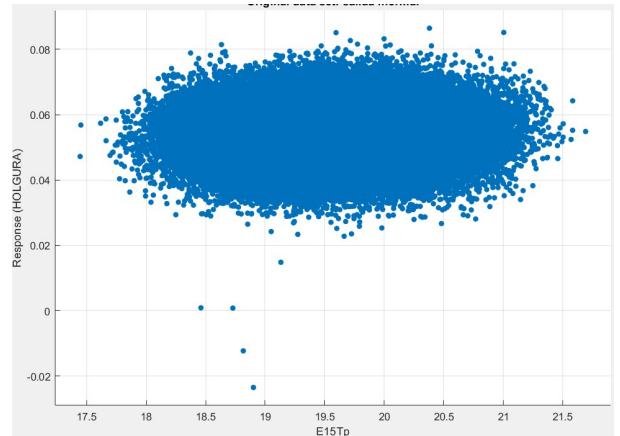
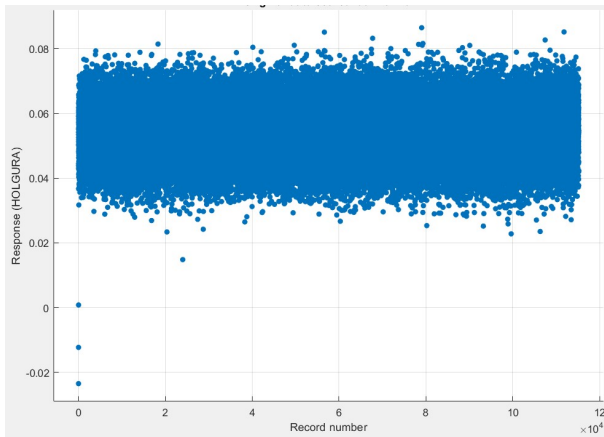
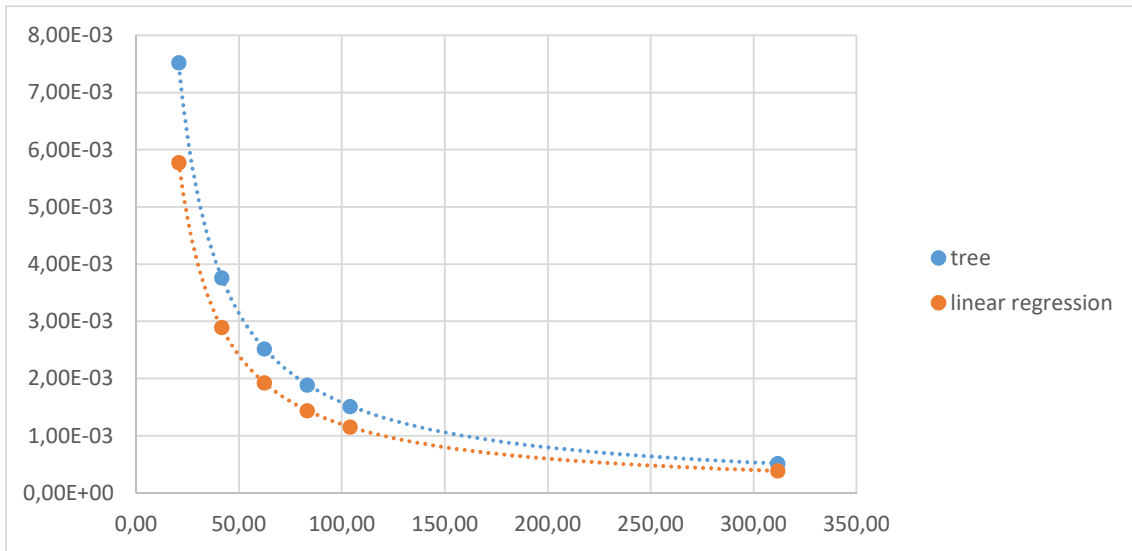


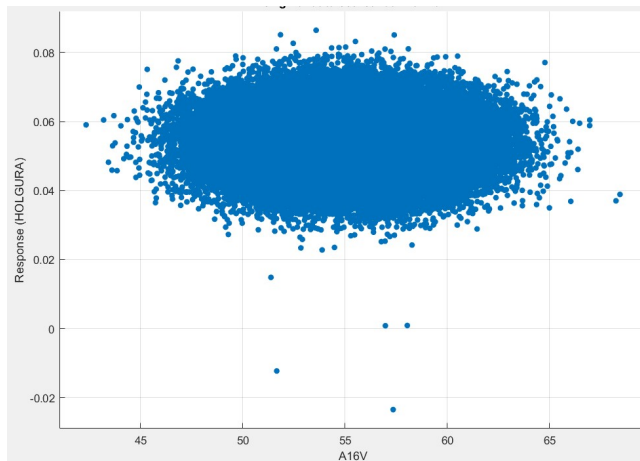
1	☆ Linear Regression	RMSE: 0.0012079
	Last change: Linear	33/33 features
2	☆ Linear Regression	RMSE: 0.0029532
	Last change: Removed 22 features	10/33 features
4	☆ Linear Regression	RMSE: 0.0012089
	Last change: Added 2 features	19/33 features



1 ☆ Linear Regression RMSE: **1.3989e-05**
Last change: Linear 11/11 features

2 ☆ Linear Regression RMSE: 1.4007e-05
Last change: Disabled PCA 5/11 features





1 ☆ Linear Regression

Last change: Linear

RMSE: **0.0012105**

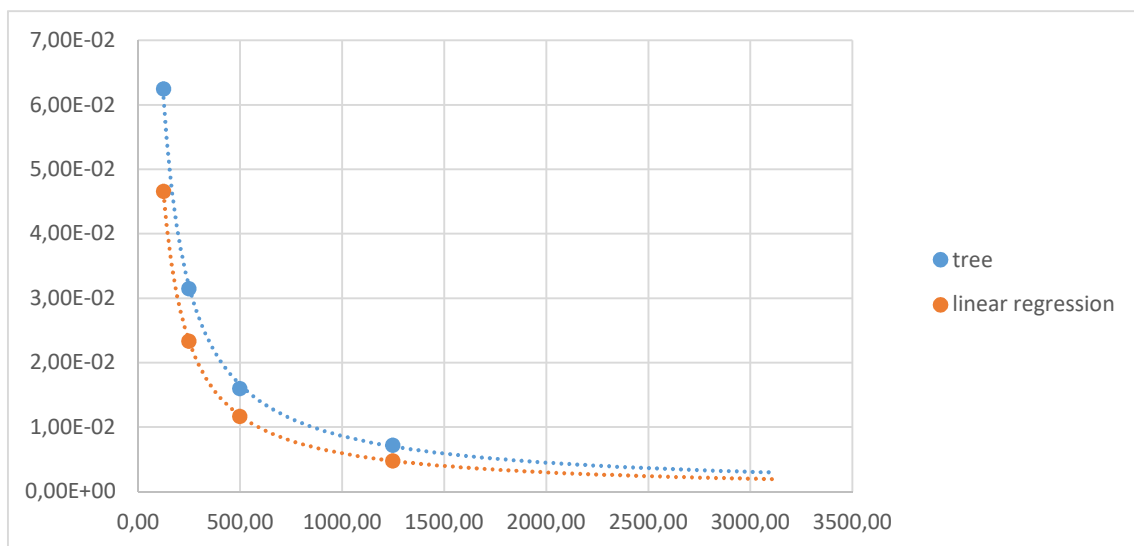
33/33 features

2 ☆ Linear Regression

Last change: Removed 13 features

RMSE: 0.0012108

19/33 features



Bibliografía

1. CORDIS: project STREAM-OD. [En línea] 18 de 3 de 2020. [Citado el: 10 de 5 de 2020.] <https://cordis.europa.eu/project/id/723082>.
2. CORDIS: project Z-Fact0r. [En línea] 6 de 4 de 2020. [Citado el: 10 de 5 de 2020.] <https://cordis.europa.eu/project/id/723906>.
3. CORDIS: project BOOST4.0. [En línea] 18 de 3 de 2020. [Citado el: 10 de 5 de 2020.] <https://cordis.europa.eu/project/id/780732>.
4. CORDIS: project TOREADOR. [En línea] 13 de 12 de 2019. [Citado el: 10 de 5 de 2020.] <https://cordis.europa.eu/project/id/688797>.